DOCUMENT RESUME

ED 050 791 LI 002 853

AUTHOR	Dolan, F. T., Ed.
TITLE	Proceedings: Second Annual Meeting, Western Canada Chapter, American Society for Information Science, September 14, 15, 1970, Vancouver Public Library.
INSTITUTION	Calgary Univ. (Alberta).
SPONS AGENCY	American Society for Information Science. Western Canada Chapter.
PUB DATE	Dec 70
NOTE	180p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Automation, Conference Reports, *Foreign Countries, *Information Retrieval, *Information Science, Information Seeking, Information Storage, *Libraries, Library Services, On Line Systems, Workshops
IDENTIFIERS	Canada, Computerized Engineering Index, Library Automation, Machine Readable Cataloging, *Selective Dissemination of Information

ABSTRACT

The two-day conference included the following: (1) a tutorial on Selective Dissemination of Information (SDI) services including those of the National Science Library, COMPENDEX, and the retrospective and current-awareness searches of "Chemical Titles;" (2) a workshop on profiling and search editing, for the aforementioned services; (3) a discussion of the experiences of those providing and those using SDI services; (4) a keynote address on on-line systems; (5) a session on MARC tapes which included search programs at the University of Alberta, a tutorial, and use of the tapes at the University of Saskatchewan, and (6) a panel discussion on joys and rewards or trials and tribulations of automating a library. A list of registrants is included. (AB)

ED050791

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PROCEEDINGS

Second Annual Meeting

Western Canada Chapter

AMERICAN SOCIETY FOR INFORMATION SCIENCE

September 14, 15, 1970

Vancouver Public Library

Information Systems
The University of Calgary
Calgary, Alberta

December, 1970

LI 002 853

PREFACE

These proceedings were compiled and edited by the Information Systems Group at the University of Calgary, with the editorial assistance of Miss K.E. Koole.

F.T. Dolan
Supervisory Editor

PROGRAM

Page

Monday, September 14

a.m.

9:00 Registration (List of Registrants) vii

9:30 Welcome and Introduction

9:45 Tutorial on SDI Services
Chairman: N. Brearley (TRIUMF, UBC) 1

9:45 - 10:00 (a) NSL, by R.J. Brongers
(Library, U of B.C.) 3

10:00 - 10:15 (b) COMPENDEX, by F.T. Dolan
(Information Services, U of Calgary) 5

10:15 - 10:30 (c) Chemical Titles and Retrospective,
by H.S. Heaps
(Computing Science, U of Alberta)

10:30 Coffee

11:00 Workshop on Profiling and Search Editing
Chairman: A. Fitzpatrick (Alberta Research Council) 11

11:00 - 11:20 (a) Chemical Titles and Retrospective,
by G.A. Cooke
(Library, Edmonton Journal) 13

11:20 - 11:40 (b) COMPENDEX, by O. Standera
(Data Centre, U of Calgary) 17

11:40 - 12:00 (c) NSL, by G. Mauerhoff
(Library, U of Sask., Saskatoon) 32

Noon Lunch

p.m.

2:00 Discussion of Experiences of Those Providing, and
Those Using, SDI Services
Chairman: K.E. Marshall (Freshwater Institute, Winnipeg) 55

3:00 Coffee

Page

p.m.

3:30 On Line Systems, by Dr. C.A. Cuadra, ASIS Distinguished Lecturer for 1970 69

Dr. Cuadra is Editor of the Annual Review of Information Science and Technology, and Manager of the Library and Documentation Systems Department at Systems Development Corporation, Santa Monica, California.

4:30 Business Meeting and Report of Officers 73

Evening Informal

Tuesday, September 15

a.m.

9:00 The MARC Tapes
Chairman: P. Simmons (School of Librarianship, U of B.C.) 81

9:00 - 9:45 (a) Search Programs for MARC Tapes at U of Alberta, by D.M. Heaps (Computing Science, U of Alberta) 83

9:45 - 10:30 (b) Tutorial on MARC Tapes, by E. Buchinsky (Library, U of Sask., Saskatoon) 95

10:30 - 11:00 (c) Use of MARC Tapes at U of Sask., by G.C. Burgis (Library, U of Sask., Saskatoon) 95

11:00 Coffee

11:30 Panel Discussion on the Joys and Rewards, or Trials and Tribulations, of Automating a Library
Chairman: T.C. Dobb (Library, SFU) 171

12:00 Closing Session, and Announcement of New Officers

LIST OF REGISTRANTS

APPELT, D.D.	Librarian, University of Saskatchewan, Saskatoon, Sask.
BAIRD, Donald A	3052 Armada, Port Coquitlam, B.C.
BERTRAN, Dr. Sheila	School of Library Science, University of Alberta Edmonton 7, Alberta
BREARLEY, Mrs. Anne	School of Librarianship, University of B.C., Vancouver 8, B.C.
BREARLEY, Neil	TRIUMF, University of British Columbia, Vancouver 8, B.C.
BRONGERS, Rein J.	Science Division, Library, University of B.C., Vancouver 8, B.C.
BRUCE, D.R.	Science Division, Library, Simon Fraser University Burnaby 2, B.C.
BUCHINSKI, Edwin J.	Library, University of Saskatchewan, Saskatoon, Sask.
BURGIS, Grover C.	Library, University of Saskatchewan, Saskatoon, Sask.
CABALLERO, Mrs. H.	6660 N.W. Marine, Vancouver 14, B.C.
CHIA, Miss Karen	Placer Development Ltd. Library, 800-1030 W. Georgia Street, Vancouver 5, B.C.
CHANDLER, Miss Beverly	General Science Dept., Cameron Library, University of Alberta, Edmonton 7, Alberta
COOKE, Mrs. G.A.	9427 75 Street, Edmonton 82, Alberta
CUADRA, Dr. Carlos A.	System Development Corporation, 2500 Colorado Avenue, Santa Monica, California 90406, U.S.A.
DENNIS, Donald J.	Systems Dept., Library, University of B.C., Vancouver 8, B.C.
DEUTSCH, Maurice	3940 East Pender Street, Apt.212, Burnaby 2, B.C.
DOBB, T.D.	Library, Simon Fraser University, Burnaby 2, B.C.
DOBBIN, Miss Gerry	University of British Columbia Library, Vancouver 8, B.C.
DOLAN, Frank T.	Data Centre, University of Calgary, Science 3B Bldg. Calgary 44, Alberta
DUNN, Ken	Systems Planning & Development, Cameron Library, University of Alberta, Edmonton 7, Alberta

ENDELMAN, M.J.	Vancouver Public Library, 750 Burrard, Vancouver, B.C.
FITZPATRICK, Art	Research Council of Alberta, Edmonton 7, Alberta
FLETCHER, Miss Mary	Research Council of Alberta, Edmonton 7, Alberta
GATZ, Mrs. Elizabeth	604 Otter Place, Richmond, B.C.
GRAHAM, Russell	3350 East Broadway, Vancouver 12, B.C.
HACKIE, Miss Sella	2419 Point Grey Road, Vancouver 9, B.C.
HEAPS, Mrs. Doreen	Dept. of Computing Science, University of Alberta Edmonton 7, Alberta
HEAPS, Stan	Dept. of Computing Science, University of Alberta Edmonton 7, Alberta
HEYWORTH, Miss Janice	Dept. of Computing Science, University of Alberta Edmonton 7, Alberta
IRWIN, Miss Louise	Canadian Agricultural Research, Vancouver, B.C.
ISAAK, David	Reference Division, McPherson Library, University of Victoria, Victoria, B.C.
JOHNSON, Mrs. Marion	Forest Products Lab., 6620 N.W. Marine Drive, Vancouver 8, B.C.
KIEHLMAN, Dr. Eberhard	Dept. of Chemistry, Simon Fraser University, Burnaby 2, B.C.
LASKOWSKI, Seno	Head, Cataloguing Division, Simon Fraser University, Burnaby 2, B.C.
LOWTHER, Mrs. Barbara	1997 Casa Marcia Crescent, Victoria, B.C.
MACDONALD, Charles	4548 East Hastings, Burnaby 2, B.C.
MACDONALD, R.W.	Library, University of British Columbia, Vancouver 8, B.C.
MACGREGOR-GREER, Mrs. Mary	Forest Research Laboratory, 506 Mansfield Avenue, Victoria, B.C.
MARSHALL, K.E.	Fisheries Research Board of Canada, Freshwater Institute, 501 University Crescent, Winnipeg 19, Manitoba.
MAUERHOFF, Georg	University of Saskatchewan Library, Saskatoon, Sask.
NEWMAN, W.L.	Computation Centre, University of Saskatchewan, Saskatoon, Sask.

PATTERSON, Mrs. A.J.	603-6th Avenue S.W., Calgary 1, Alberta
PIERCE-JONES, David	Associate Librarian, University of Manitoba Library Winnipeg 19, Manitoba.
ROBERTS, John J.	Whiteshell Nuclear Research Establishement, Atomic Energy of Canada Ltd., Finawa, Manitoba.
RODNEY, Mrs. Helen	McPherson Library, University of Victoria, Victoria, B.C.
SCHICK, Miss Patricia	General Science, Cameron Library, University of Alberta, Edmonton 7, Alberta
SCHLAUCH, Edward	118-7 Avenue South, Lethbridge, Alberta
SIM, Derek	Sherritt Gordon Mines Ltd., Fort Saskatchewan, Alberta
SIMMONS, Peter	School of Librarianship, University of B.C., Vancouver 8, B.C.
SPAULDING, J. Gordon	Dept. of English, University of B.C., Vancouver 8, B.C.
STANDERA, Olderich	University of Calgary, Calgary 44, Alberta
THOMAS, Lawrence E.	128 N. Stratford, Burnaby B.C.
THOMPSON. Gordon R.	6612-94 Avenue, Edmonton 83, Alberta
TURNER, Mrs. Ann	#807-1825 Haro Street, Vancouver 5, B.C.
WILIMOVSKY, Miss Diana	MacMillan Bloedel Research Ltd., 3350 E. Broadway Vancouver 12, B.C.
WILLIAMS, Miss Lynda	Ste. 3-25 Longside Street, Winnipeg 1, Manitoba
YOUNG, B.S.	Sherritt Gordon Mines Ltd., Fort Saskatchewan, Alberta
YOUNG, Miss Adwena	28 Pine Street, Namaimo, B.C.

TUTORIAL ON SDI SERVICES

N. Brearley
Chairman

Tutorial on SDI Services

Chairman's Summary

by N. Brearley, TRIUMF Project
University of British Columbia

The first session of the morning consisted of a Tutorial on SDI Services chaired by Neil Brearley (TRIUMF, University of British Columbia). The panelists were: Rein J. Brongers (Science Division, Library, University of British Columbia), Frank T. Dolan (Information Services, University of Calgary) and H. Stanley Heaps (Department of Computing Science, University of Alberta).

Rein Brongers gave a short talk on the CAN/SDI service offered by the National Science Library. This system became a nation-wide public service in the spring of 1969. Its data base now consists of the following tape services: Chemical Titles, Chemical Abstracts Condensates, ISIS (Institute for Scientific Information) source tapes, citation tapes, and organization tapes, and INSPEC (Information Service in Physics, Electrotechnology and Control).

He gave a brief description of each of these, stressing differences in coverage and searchability, and wound up with a plea for more and better "sales promotion" and user education.

Frank Dolan described the Compendex data base which is now available for searching at the University of Calgary. This is the tape version of Engineering Index Monthly. A profiling guide has been written by Oldrich Standera and is available from Information Systems, University of Calgary.

Service to persons other than those affiliated with the University of Calgary is available through the Alberta Information Retrieval Association.

The Compendex tapes began in January, 1969, and this sets the limit on retrospective searches.

The cost of the service is $100 per year for 40 terms, each additional 10 terms costing $20. Charges will be based on the average number of terms used during the year.

Retrospective and current-awareness searches of Chemical Titles tapes were described by Stan Heaps. This service is available through the Alberta Information Retrieval Association. Current searches duplicate the service offered by the National Science Library, but AIRA claims that their service can be cheaper for the smaller profiles (<40 terms).

Overview of COMPENDEX

by F.T. Dolan, Mgr. Information Systems,
University of Calgary

The acronym COMPENDEX stands for Computerized Engineering Index.

Engineering Index Incorporated in New York city markets this package in three parts:

COMPENDEX $6800/year	Monthly Magnetic Tape Service	$6300/year
	Monthly Indexes	$400/year
	Annual Index	$200/year

We use an IBM system called TEXT-PAC to process this data base. This free-text processor was obtained at no cost from IBM.

The Engineering Index data base is compiled from more than 3500 sources which include periodicals, books, technical reports and proceedings. It includes approximately 5000 records per month.

Each record in this data base contains the title, personal and corporate author(s), complete bibliographical information, Ei# and full informative abstract. Each of the data elements is labelled so that the computer knows one from the other for controlled search.

TEXT-PAC allows us to search the full text of this data base in two modes: current awareness and retrospective search.

In current awareness, we try to capture an engineer's present interest using keywords and logical and syntactical connectors. (Oldrich will give you more detail on how this is done.) We then compare a batch of such profiles against each record in the data base.

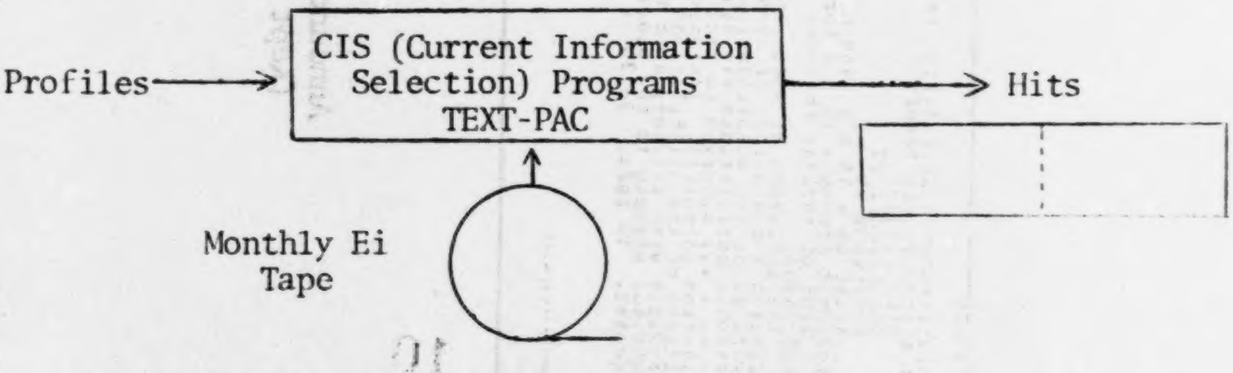
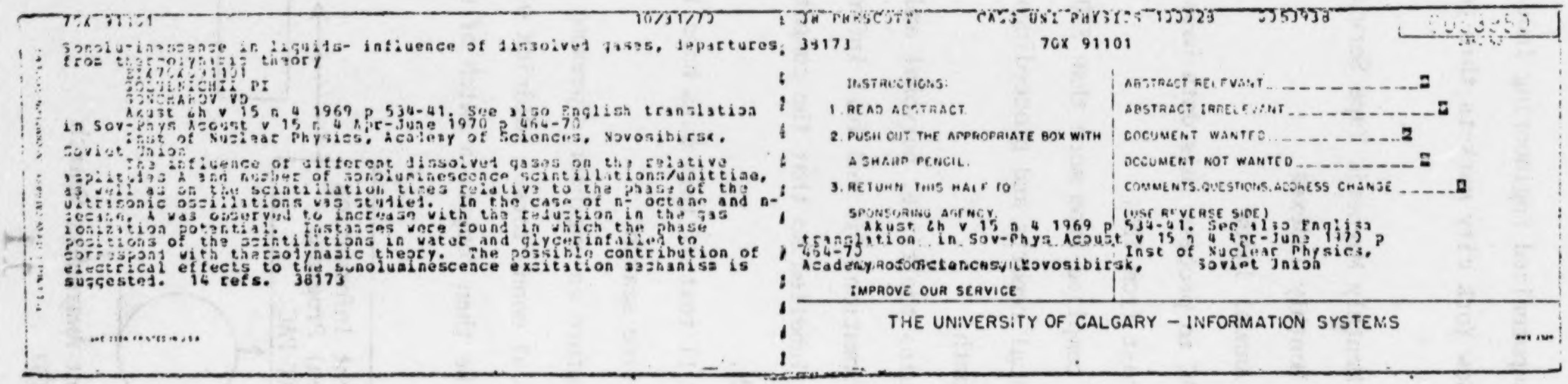


Fig. 1 Current Awareness Search

Each user gets his hits in double card format as shown in Fig. 2.

Note TEXT-PAC's ability to handle upper and lower case.



70X 91101 10/31/70 JW PRESCOTT CAGS UNI PHYSICS 100028 0050938

Sonoluminescence in liquids- influence of dissolved gases, departures from thermodynamic theory 38173

[illegible]
GOLUBNICHII PI
GONCHAROV VD
Akust Zh v 15 n 4 1969 p 534-41. See also English translation in Sov-Phys Acoust v 15 n 4 Apr-June 1970 p 464-70
Inst of Nuclear Physics, Academy of Sciences, Novosibirsk, Soviet Union

The influence of different dissolved gases on the relative amplitudes A and number of sonoluminescence scintillations/unittime, as well as on the scintillation times relative to the phase of the ultrasonic oscillations was studied. In the case of n- octane and n-decane, A was observed to increase with the reduction in the gas ionization potential. Instances were found in which the phase positions of the scintillations in water and glycerinfailed to correspond with thermodynamic theory. The possible contribution of electrical effects to the sonoluminescence excitation mechanism is suggested. 14 refs. 38173

70X 91101

INSTRUCTIONS:

1. READ ABSTRACT.
2. PUSH OUT THE APPROPRIATE BOX WITH A SHARP PENCIL.
3. RETURN THIS HALF TO SPONSORING AGENCY.

IMPROVE OUR SERVICE

ABSTRACT RELEVANT ☐
ABSTRACT IRRELEVANT ☐
DOCUMENT WANTED ☐
DOCUMENT NOT WANTED ☐
COMMENTS, QUESTIONS, ADDRESS CHANGE ☐
(USE REVERSE SIDE)

Akust Zh v 15 n 4 1969 p 534-41. See also English translation in Sov-Phys Acoust v 15 n 4 Apr-June 1970 p 464-70 Inst of Nuclear Physics, Academy of Sciences, Novosibirsk, Soviet Union

THE UNIVERSITY OF CALGARY – INFORMATION SYSTEMS

Announcement Half
(Kept by User)

Response Half
(Returned to Us)

Fig. 2 Current Awareness Hit (On Cards)

The Retrospective Search mode is very similar.

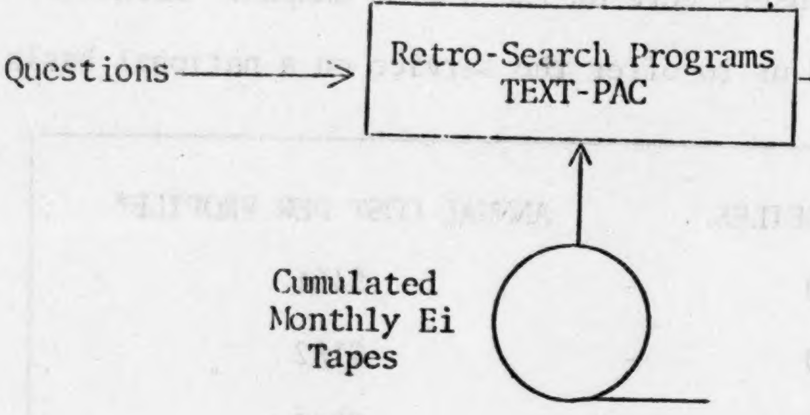
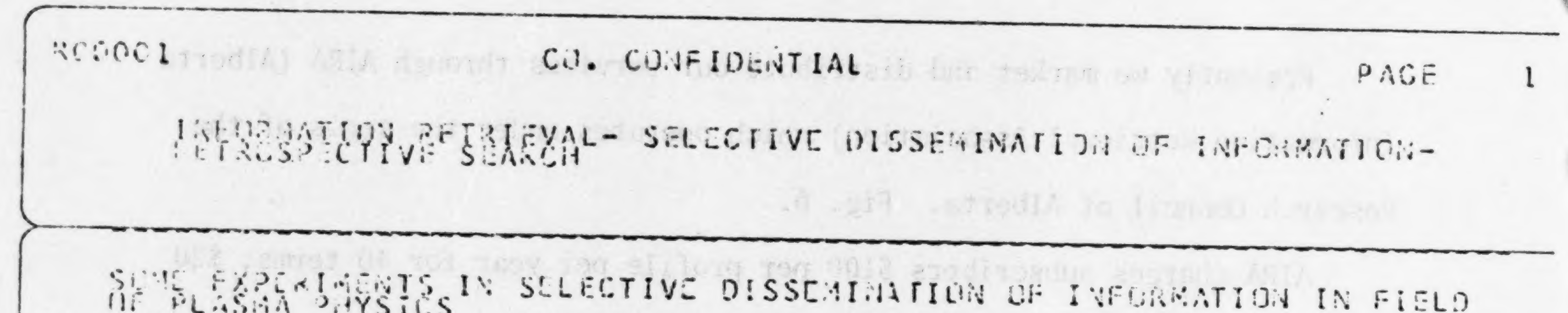


Fig. 3 Retrospective Search



```
RC0001            CO. CONFIDENTIAL                        PAGE      1

     INFORMATION RETRIEVAL- SELECTIVE DISSEMINATION OF INFORMATION-
     RETROSPECTIVE SEARCH

  SOME EXPERIMENTS IN SELECTIVE DISSEMINATION OF INFORMATION IN FIELD
  OF PLASMA PHYSICS
  INFORMATION STORAGE AND RETRIEVAL 30274

        EIX69X100165

        ANTHONY LJ
        CHENEY, AG
        WHELAN EK

        30274

         INFORMATION DSTORAGE & RETRIEVAL V 4 N 2 JUNE 1968 P 187- 20

        SMALL- SCALE, COMPUTER- BASED SDI SYSTEM IN PLASMA PHYSICS AND
  RELATED SUBJECTS USES TITLE INPUT ONLY IN ORDER TO MINIMIZE INPUT
  COSTS;  IMPLICATIONS OF THIS APPROACH AND ITS EFFECT UPON SYSTEM
  PARAMETERS IS DISCUSSED;  SOME COMPARISON OF COSTS OF COMPUTERBASED
  SYSTEM WITH THOSE OF MANUAL SYSTEM IS MADE;  FURTHER EXPERIMENTS ARE
  DESCRIBED IN WHICH SERVICE IS EXPANDED TO EXTERNAL USERS ON WORLD-
  WIDE BASIS;  SOME OBSERVATIONS ON FUTURE DEVELOPMENT AND ORGANIZATION
  OF COMPUTER- ASSISTED SERVICES, THEIR POSSIBILITIES AND MAIN PROBLEMS
  WHICH ARE LIKELY TO ARISE. 9 REFS.  30274

        INFORMATION STORAGE AND RETRIEVAL
        00-A157
        00-A065
        00-A155
        00-A300
        PLASMAS
        COMPUTERS
```

Fig. 4 Retrospective Search Hit (On Paper)

When Information Systems first began running the COMPENDEX service in late fall '69, all our users were on the U of C Campus. Economies of scale, Fig. 5, soon forced us to offer the service on a national basis.

# OF PROFILES		ANNUAL COST PER PROFILE*
70	-	$454
210	-	$182
280	-	$143

*All Costs: Salaries, Data Base, Computer Time, Overhead, etc. are incorporated here.

Fig. 5 Economies of Scale

Presently we market and distribute our services through AIRA (Alberta Information Retrieval Association) which operates under the aegis of the Research Council of Alberta. Fig. 6.

AIRA charges subscribers $100 per profile per year for 40 terms; $20 per profile per year for each additional 10 terms.

COMPENDEX has grown steadily during its first year of operation and we anticipate in the near future, that this service will be paying for itself.

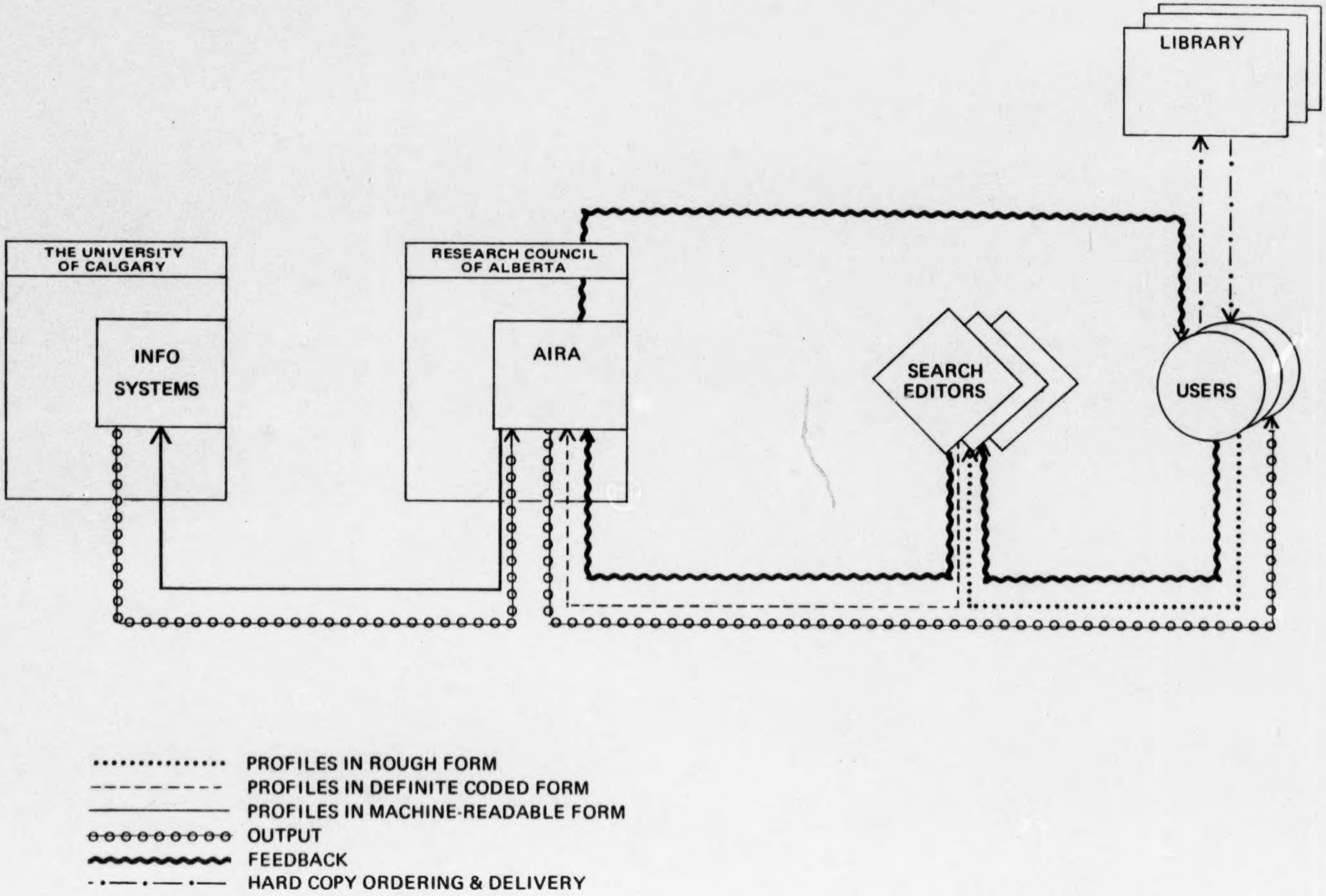

THE UNIVERSITY OF CALGARY
INFO SYSTEMS
RESEARCH COUNCIL OF ALBERTA
AIRA
SEARCH EDITORS
USERS
LIBRARY
PROFILES IN ROUGH FORM
PROFILES IN DEFINITE CODED FORM
PROFILES IN MACHINE-READABLE FORM
OUTPUT
FEEDBACK
HARD COPY ORDERING & DELIVERY

Fig. 6 COMPENDEX Flow

Workshop on Profiling and Search Editing

A. Fitzpatrick
Chairman

Chemical Titles and Retrospective

by G. A. Cooke, Library,
Edmonton Journal

Most of what I have to say is very ably put forth in the two guides from the Alberta Information Retrieval Association entitled "AIRA/CT Profile Design" and "User's Guide for Retrospective Searches of Chemical Titles".

Therefore, what I would like to try to do to-day is to pinpoint some of the highlights from these publications and to pass along a few tips learned through trial and error on my part. I will first use Chemical Titles tapes as my example and will then add some remarks regarding C.T. Retrospective searches at the end.

I realize that many of you are familiar with the terms I will use and hope you will bear with me while I briefly go over them for the benefit of those who are not.

In order to search any data base one must construct a search PROFILE. This profile consists of carefully selected words or phrases which will best describe your (the User's) needs. One must always take into account the vocabulary of the data base and particularly with C.T. tapes, one must keep in mind that only titles are being searched besides authors and journal codens.

The words or phrases of a profile can be interconnected by the use of the Boolean logic operators AND, OR and NOT. A group of terms connected by OR is called a PARAMETER (or a concept). Parameters can be grouped by the use of AND. NOT is used to exclude certain unwanted terms and can be particularly useful for eliminating certain journals or authors from the output.

To take a very simple example:

You are interested in obtaining notification of all articles on the subject of cats and/or dogs. A simple one parameter profile for this will

suffice viz. DOGS
or CATS

Meaning that any title containing the word dogs or the word cats will fill the bill.

But you may only be interested in papers dealing with the interaction of dogs and cats. Then the profile would become two parameters,viz. DOGS AND CATS meaning that the title must contain both words to fill the bill -- or score a hit as we say.

Now let us say that you are not interested in the article if the cat is a persian cat. Then the profile becomes DOGS AND CATS NOT PERSIAN. This form of profile would also exclude persian dogs, or even an article on dogs and cats in the Persian Gulf. There are many ways using these logic operators (AND, OR and NOT) that single words or groups of words can be combined to specify the articles of interest.

All of this is dependent, of course, on such words occurring in the title. Thus for C.T. searches, I strongly recommend you keep the search profile simple -- at the most two parameters. Remember a parameter is a group of words or phrases joined by OR logic. Thus any one of a string of say 20 possible words will score a hit.

I think you will be beginning to see that to list all possible variation of terms, particularly for singular, plurals and verb endings, is both time consuming and costly since most services charge you per term (or per a given number of terms). So we resort to the technique of TRUNCATION. This is simply a chopping off of letters at the beginning or end of a word so that a stem (or root) which covers several variations is left. The point of truncation is often denoted by a *.

Thus DOGS* will not only retrieve dog, but dogs, dogged, dogeral, dogwood, etc. Sometimes, as you can see if you use a dictionary to see what could be retrieved by CAT* (catastrophic), it is best to forgo the economy of truncation and to spell out the individual words. It is up to you, depending on how many irrelevant hits (or NOISE) you are prepared to support.

CT Search programs, developed at the University of Alberta and used by AIRA, permit truncation at the beginning or end of a word. Front truncation can be very useful to some, eg. an organic chemist wishing to cover a family of compounds, but it can also produce some highly irrelevant and very unexpected hits. Front truncation also markedly increases search times, a point worth remembering if you pay for searches by computer time used.

The search profiles for C.T. tapes also permit the use of weights. Values from -999 to +999 can be attributed to a search term. As the tapes are searched, a count is made of the total value depending on the weight ascribed to the terms which are causing a hit. A certain threshhold weight, which must be equalled or exceeded, is specified. The proper use of weights is best learned with practise and is not recommended for searching titles. However, one use that I have found particularly helpful is for ordering the output from a profile which is used to search for different aspects of a subject. For example, one user of your group may be interested in the use of clays, another in the chemical composition and another in its extraction, a separation will be obtained in the print-out, since all hits are printed in descending order of weight. A threshhold weight of 1 is given to such a profile so that no hits are missed. You will note I have progressed in numbers by a multiple of 4. This usually gives sufficient separation in a search of titles since it will require four words from one group to score in order to be confused with one hit from the next group. Actually, a simple progression of 1, 2, 4, 8, 16, 32, --- works very well with titles. This method also helps for quick reviewing of a profile, since one can soon recognize those terms which are giving valuable hits.

To turn briefly to C.T. Retrospective searches. Search programs for this were also developed by the Computing Science Department of the University of Alberta. Profile preparation is essentially the same though there are some important differences:

1. There is no weighting of terms.
2. Certain common connective words (e.g. A, AND, BY, -- -) are excluded as search terms.
3. The output includes a listing of all words that will be retrieved by any truncated terms in the profile.
4. There are no search capabilities as yet for Authors.

The point to remember about these tapes is that titles in the early years were not very descriptive, making the formulation of an adequate search profile very difficult.

A very much more extensive search program is being developed which will allow searches on authors and the use of extended logic operators.

I think that is all; just a final reminder -- remember these services are only searching titles so keep the profiles simple. Also, formulate the profiles on a broad basis at first until you learn by experience which terms retrieve the best for you.

COMPENDEX PROFILING

by O. R. Standera, Information Systems,
The University of Calgary

0. Abstract The present paper is designed for search editors, users and all those interested in how profiles are constructed in the TEXT-PAC system adopted for the COMPENDEX service at The University of Calgary. After explaining the terminology used, the author indicates how the searching function of words may be modified by truncation and capitalization. Logical connectors are defined and their use in the three levels of back-referencing is illustrated. Practical examples show how to formulate a simple information need into a profile. This paper is an essence of the COMPENDEX PROFILING GUIDE.

1. INTRODUCTION

This morning, F. T. Dolan, Manager, Information Systems, The University of Calgary, has given you an overview of the COMPENDEX service and the TEXT-PAC system. My task is to complete this by giving more details about the profiling technique. I will not be dealing with the profile form which is amply described in our yellow COMPENDEX PROFILING GUIDE. This guide was distributed to all of you.

The present paper and the COMPENDEX Profiling Guide provide introductory information about profile set-up. We intend, however, to publish a brief "philosophy" of profiling which would enable each search editor to adjust the running profiles to the desired level of performance. This idea was substantiated by the views of some attendees of this meeting as well as other users.

We hope to make it available shortly.

2. TERMINOLOGY

First of all, let me say a few words about the terminology, which may be more ambiguous in the field of search editing than anywhere else.

The basic elements of a profile are profile words. Profile words (terms) are connected to each other by means of logical connectors forming the concepts which constitute search expressions. One or more search expressions form a question (profile).

Profile words (terms), concepts,or search expressions may be represented by logical symbols. Notice that the search expressions are denoted by CON in the COMPENDEX Profile Submission Forms. (The original TEXT-PAC documentation uses "concepts" where we introduced "search expressions.")

The following example (Fig. 1) is designed to clear up terminology:

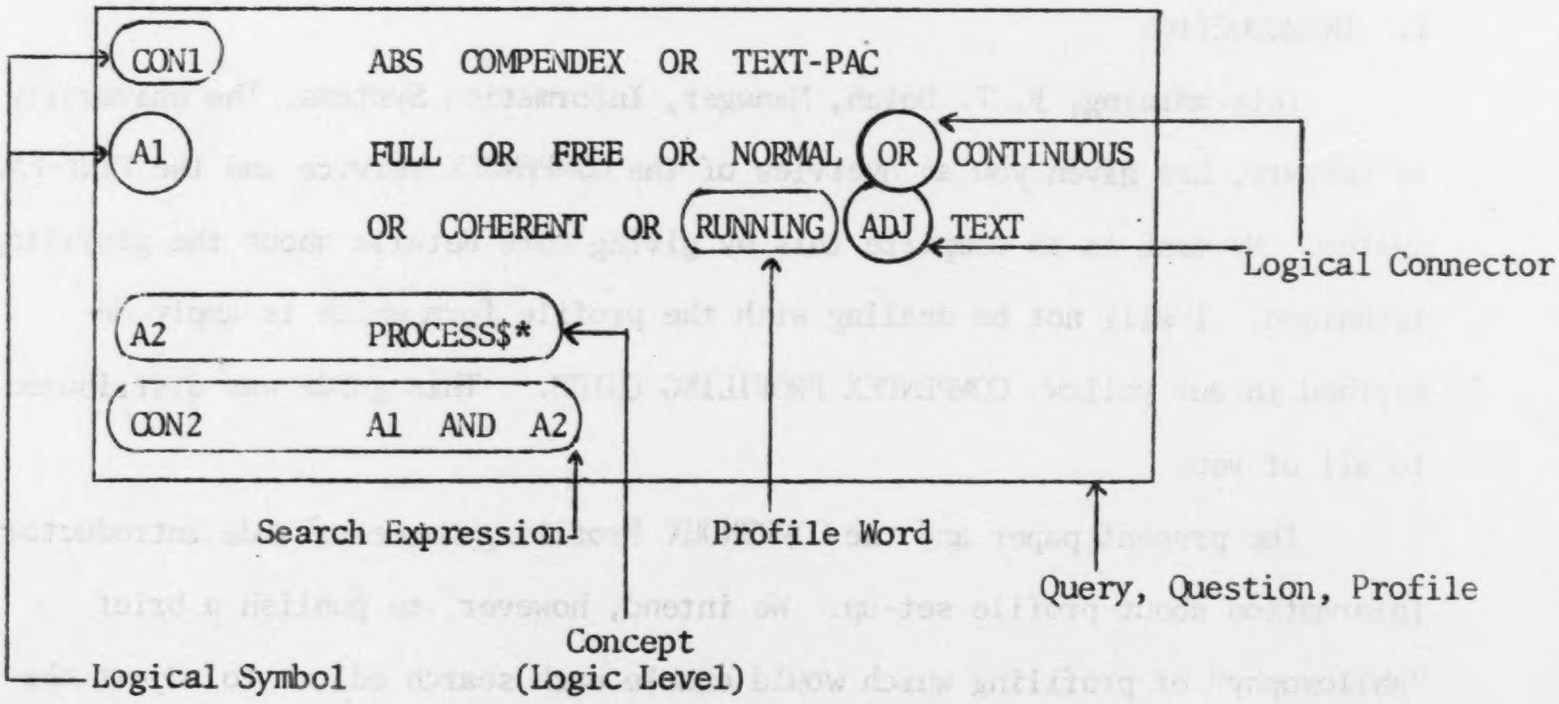

CON1 ABS COMPENDEX OR TEXT-PAC
A1 FULL OR FREE OR NORMAL OR CONTINUOUS
OR COHERENT OR RUNNING ADJ TEXT
A2 PROCESS$*
CON2 A1 AND A2
Logical Connector
Search Expression
Profile Word
Query, Question, Profile
Concept
(Logic Level)
Logical Symbol

Figure 1

Terminology used in The University of Calgary COMPENDEX service

2.1 WORDS

Having established our terminology, let us have a close look at the fundamental profile element--a word.

2.1.1 Length of words, spacing. The length of any profile word may be up to 38 characters. However, only the first 20 characters are searched.

Leave a space after every word and after every logical connector. Two words must always be separated by any of the logical connectors.

2.1.2 Truncation. It is sometimes desirable to search on word stems rather than on the full words.

TEXT-PAC allows right end truncation only. Truncation can be done in two ways:

Selective truncation may extend as far as six characters past the root. ORGANI$$$$$$ will cover ORGANIZE, ORGANIZER, ORGANIZERS, as well as ORGANIZING, ORGANIZATION. As we may use only six dollar signs, we have to use unconditional truncation if we also want ORGANIZATIONAL to be included in our profile formulation.

Unconditional truncation ORGANI$* will cover all possible endings of the given root as far as twenty characters. The root may consist of a minimum of one character.

When using this profiling facility one must carefully consider all possible words that might be matched. One might save several seconds by indiscriminate truncation but lose a considerable period of time getting through irrelevant information produced.

For example if you are interested in programming and programs of retrieval systems, specifying PROGR$* would find not only desired programming and programs, but also unwanted progress, progression, etc.

2.1.3 Capitalization. Because TEXT-PAC handles both lower case and upper case printing, you may refine your profile even more by taking advantage of the following rules. It should be pointed out that capitalization is not widely used though there are specific cases where it is warranted.

Assuming you have not specified capitalization, the profile word will be a match if there is such a word in the data base, no matter if the letters are in upper case, lower case, or in any combination.

"GIPSY" in the profile formulation will match any information about gipsy as well as about the acronym GIPSY denoting an information system.

If you specify one "at sign," it will match only all upper case characters (GIPSY) or initial capitalization (Gipsy).

Correct specification: @ GIPSY

Two "at signs" (@@GIPSY) will find only all upper case characters.

If you wish to have hits only with a word containing mixed upper and lower case letters, then you may use the number sign. #PH will match only pH which means the concentration of hydrogen ions.

As you may have recognized you can make your job a lot easier without specifying the capitalization unless it is necessary to do so.

3. LOGICAL CONNECTORS

Logical connectors are used to connect individual profile words or logical symbols to build more complex units: concepts and search expressions. In TEXT-PAC we use the following logical connectors:

OR

AND

WITH

ADJ

NOT

ABSOLUTE

CONTROL

NOT-CONTROL

OR: Logical connector OR combines profile words or logical symbols indicating that any of them will satisfy the user's requirement. In our example we are interested in "text" which may be specified as:

FULL OR FREE OR NORMAL OR CONTINUOUS OR COHERENT OR RUNNING . . . .

AND: Logical connector AND identifies the profile words or concepts which must be present jointly in a data base record for the hit to occur. A maximum of 15 profile words may be connected by AND.

For example, "USERS' AND FEEDBACK" means that the hit will only result if both of these words occur in the same document. It is evident that we might get some irrelevant hits if one sentence dealt generally with "USERS' REACTION" and another described "FEEDBACK" in electronics.

ADJ, WITH: Two profile words or concepts linked by ADJ must occur in the order specified to bring about a match.

SEARCH ADJ EDIT$$$

The logical connector WITH will cause a hit if the connected profile words or concepts are found in the same sentence of the document.

RETROSPECTIVE WITH SEARCH$$$ WITH STORAGE OR CORE . . . .

will produce a hit in any of the following contexts:

STORAGE REQUIRED BY RETROSPECTIVE SEARCHING . . . .

RETROSPECTIVE SEARCH NEEDS MORE STORAGE THAN . . . .

STORAGE CONSIDERATIONS FOR CURRENT AWARENESS, RETROSPECTIVE SEARCHING . . . .

Concerning the use of ADJ or WITH jointly with CONTROL and NOT-CONTROL logic, see the paragraph on CONTROL.

After you have formulated a few profiles in TEXT-PAC system, you will appreciate the way you can make your concepts and search expressions broader or narrower, thus obtaining more or less hits.

AND ↑
WITH more hits, less relevance
ADJ

This arrow shows the direction of obtaining more hits, although you may get more irrelevant information at the same time.

Remember two rules for proper use of ADJ or WITH:

(1) Only one type of logical connector may occur in a concept or search expression. There is one exception: you can use OR logic inside ADJ or WITH logic provided you connect profile words and not logical symbols denoting concepts.

__RIGHT__

INFORMATION OR RETRIEVAL ADJ SYSTEM$ OR CENTER$ . . . .

PROFILE$ OR QUESTION$ OR QUER$$$ WITH CONSTRUCT$$$ OR SET$$$$ . . . .

__WRONG__

A1 or A2 WITH A13

(2) Using ADJ or WITH logical connectors to connect two or more logical symbols which denote concepts, always make sure that the logical symbols cited represent words joined by OR logic (another formulation of the above example):

RIGHT

A1 INFORMATION OR RETRIEVAL

A2 SYSTEM$ OR CENTER$ OR . . . .

A3 A1 WITH A2

ALSO: A3 A1 ADJ A2

WRONG

A1 INFORMATION ADJ RETRIEVAL

A2 SYSTEM$ OR CENTER$

A3 PROFILE$ OR QUESTION$ OR QUER$$$ WITH CONSTRUCT$$$ OR SET$*

CON1 A1 WITH A2 WITH A3

The proper way to formulate a search expression such as this would be

CON1 A1 AND A2 AND A3

ABS: The logical connector for the ABSOLUTE logic is identified as ABS. When using ABS the hit will result with occurrence of any word accompanied by ABS in any context whatsoever. Remember that ABS may be used only in search expressions (not in concepts) and must be the first word of logic data. In this case any document containing the profile words COMPENDEX or TEXT-PAC will be quoted as a hit regardless of all the other logic.

CON1 ABS COMPENDEX OR TEXT-PAC

NOT: The NOT logical connector denotes the profile words which we do not wish to cause a hit. It overrides any other logical connector except ABS. This implies that if a given document contains a profile word which was denoted by NOT and another profile word specified by ABS, this document will become a hit. Keep in mind that you can only use NOT in search expressions and it must be the first word of logic data, e.g., the user wants all the information specified but he has enough information dealing with "libraries" already available and desires it to be excluded:

CON 7 NOT LIBRAR$$$$

GENERAL REMARKS ON QUESTION FORMULATION. It should be noted that matching profile against data base is done against the search expressions.

When constructing your profile, remember to include terms which are synonymous or closely related to your basic terms. Then formulate as many search expressions as needed to cover your information request.

Label the concepts with logical symbols A1, A2, A3 and so on. Label the search expressions CON 1, CON 2, CON 3 . . . .

Any concept may contain either logical symbols or words but not both together.

e.g. CON 1 A1 AND A2
CON 9 A6 OR A7 OR A8
A8 CURRENT ADJ AWARENESS

WRONG

A21 ECONOMICS AND A20

The same rule applies to search expressions. Remember you may use only one type of logical connector in any one logic level (concept or search expression). The only exception is mentioned in the section dealing with ADJ and WITH.

CONTROL and NOT-CONTROL: These features of the TEXT-PAC logic are logical connectors only in a broader sense, however, since they modify the function of logical connectors, they were included here. In order to understand their purpose, we must familiarize ourselves with the format of the COMPENDEX record. Any data-base record encompasses all or most of the following data elements called "Print Controls."

The numbers on the left side denote the "Print Controls." TEXT-PAC enables the user to search any or all of these elements:

00	Title
09	Subject heading (subheading may also be present), EI number
10	ID (identification number) which is the internally assigned sequential number
201	Author (as many as 99 authors may be specified under 201-299)
3	EI number
4 Z	Citation (Source)
401	Author affiliated (or first author if more than one specified)
50	Abstract
60	Subject heading (and subheading)
610	(to 649) Sales codes relating to the former Card Service of EI. These will soon be replaced by CAL identifying areas in the CARD-A-LERT service of EI.
650	Access words or keywords

The outstanding feature of the TEXT-PAC system is its ability to search the entire record as we have shown. Searching limited to one or more of these print controls is possible, although not typical. It may only be justified for example if we need all papers published by an author. Then we search only in the print control 2$$ e.g., CON14 WHITBY CONTROL2$$ ADJ DK.

The means for conducting a search in this way is called "CONTROL" logic. If we use "CONTROL" then the hit will only be achieved if the logic specified in the profile matches the logic in the specified print control of any data base record.

The rules governing use of CONTROL logic are:

1. We can use it only with profile words (not logical symbols)
2. The CONTROL is followed by print control without blank
3. As many as seven print controls may follow a word. They are separated by commas without blanks
4. Print controls are listed in ascending order
5. Print controls may be masked by dollar signs on the second and third character
6. When using ADJ or WITH logical connectors, you may only use CONTROL logic with the first profile word to the left of the first ADJ or WITH.

The "NOT-CONTROL" logic is subject generally to the same regulations. It is used if we do not want the search to be conducted in a certain print control e.g. we do not need our own papers because we have them thoroughly documented.

CON15 STANDERA NOT-CONTROL2$$ ADJ OR

Keep in mind that limiting the search uses only partially the capabilities of the system.

4. THREE LEVELS OF BACK-REFERENCING

When constructing your search expressions you may use the concepts and search expressions in three levels. This is an excellent feature of TEXT-PAC and the following figure (Fig. 2) will clear up the principles involved.

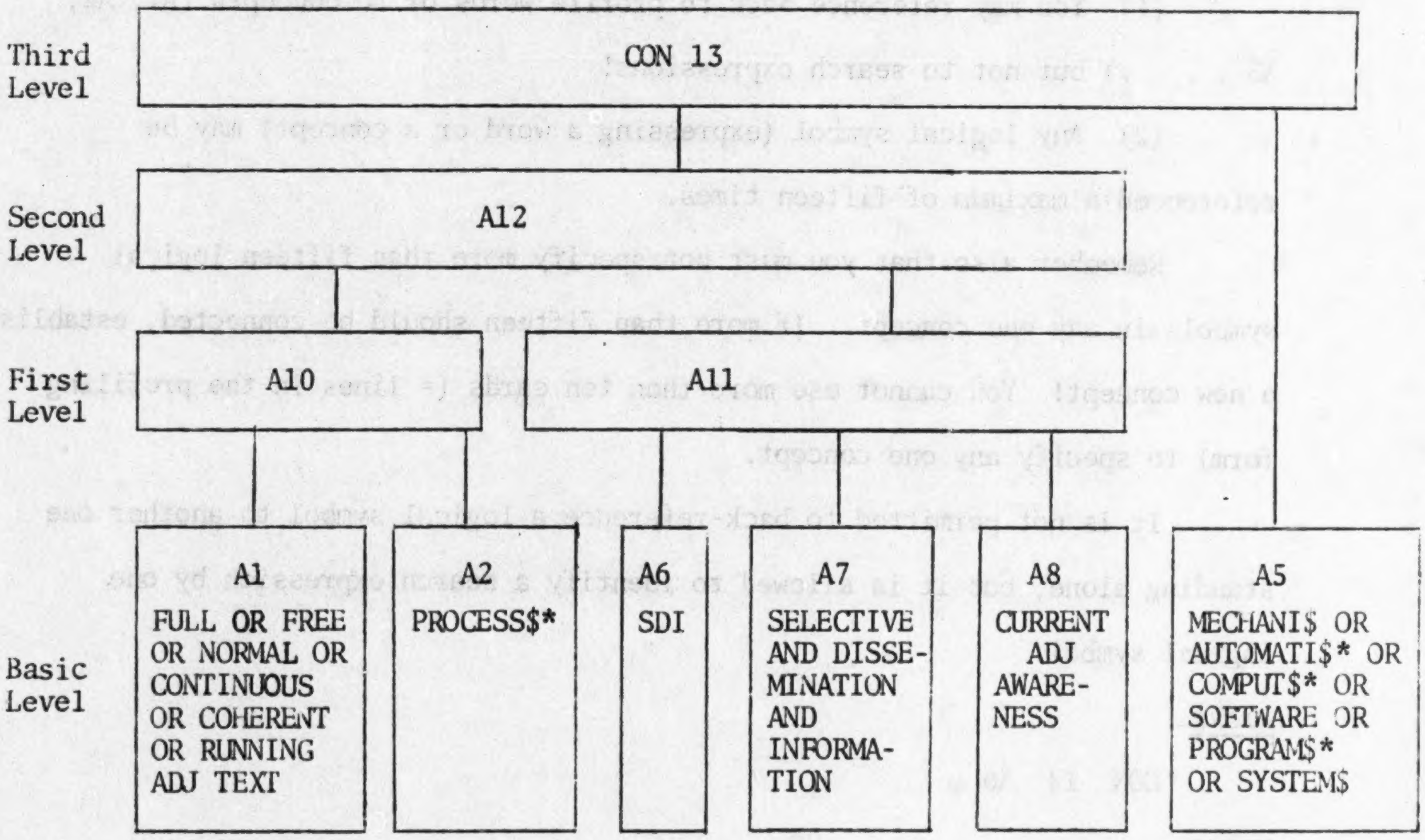
You will notice on the following figure that you may reference e.g.:

(1) the search expression back to A12

(2) A12 back to A10 and A11

(3) A10 back to A1 and A2

More levels of referencing will cause an error reported by an error message of the computer.


Third Level
CON 13
Second Level
A12
First Level
A10
A11
Basic Level
A1
FULL OR FREE OR NORMAL OR CONTINUOUS OR COHERENT OR RUNNING ADJ TEXT
A2
PROCESS$*
A6
SDI
A7
SELECTIVE AND DISSEMINATION AND INFORMATION
A8
CURRENT ADJ AWARENESS
A5
MECHANI$ OR AUTOMATI$* OR COMPUT$* OR SOFTWARE OR PROGRAM$* OR SYSTEM$

A1	FULL OR FREE OR NORMAL OR CONTINUOUS OR COHERENT OR RUNNING ADJ TEXT
A2	PROCESS$*
A5	MECHANI$* OR AUTOMATI$* OR COMPUT$* OR SOFTWARE OR PROGRAM$* OR SYSTEM$
A6	SDI
A7	SELECTIVE AND DISSEMINATION AND INFORMATION
A8	CURRENT ADJ AWARENESS
A10	A1 AND A2
A11	A6 OR A7 OR A8
A12	A10 AND A11
CON 13	A5 AND A12

Figure 2

Levels of Back-referencing

There are two more rules concerning back-referencing in the three-level structure:

(1) You may reference back to profile words or to concepts (A1, A2, A3 . . . .) but not to search expressions!

(2) Any logical symbol (expressing a word or a concept) may be referenced a maximum of fifteen times.

Remember also that you must not specify more than fifteen logical symbols in any one concept. If more than fifteen should be connected, establish a new concept! You cannot use more than ten cards (= lines in the profiling form) to specify any one concept.

It is not permitted to back-reference a logical symbol to another one standing alone, but it is allowed to identify a search expression by one logical symbol:

RIGHT

CON 14 A6

WRONG

A13 A6

5. EXAMPLES

We have familiarized ourselves with the basic principles governing the profile generation. Let us now show the practical implications of these rules on a few simple profiles.

Most of your profiles will have a simple structure. We recommend a straightforward simple structure as it is easy to establish and maintain. Rather than one complicated concept use two simpler ones. The same applies to search expressions. After some time you will find it easy to set up profiles of any degree of sophistication shown below. A few examples of simplified profiles are:

(1) Narrative statement. I need information pertaining to synthetic (plastic) foam, as far as it is related to the manufacture. Also properties of synthetic foam are of interest.

Profile

A1 SYNTHETIC OR PLASTIC

A2 FOAM$

A3 A1 WITH A2

A4 PROPERT$$$ OR CHARACTERISTIC$ OR MANUFACTUR$$$ OR PRODUC$$$$

CON1 A3 AND A4

Explanation

Dollar signs e.g. in PRODUC$$$$ mean that this formulation covers "PRODUCTION," "PRODUCE," "PRODUCER," etc. "FOAM$" covers both "FOAM" and "FOAMS." A3 connects "FOAM$" with either "SYNTHETIC" or "PLASTIC." "WITH" implies occurrence of both A1 and A2 in the same sentence. CON1 links A3 and any of the terms under A4. Terms or symbols connected by "AND" must occur in the same record to produce a hit.

(2) Narrative Statement. The same as under (1).

Profile

A1 SYNTHETIC ADJ FOAM$

A2 PLASTIC ADJ FOAM$

A3 A1 OR A2

A4 PROPERT$$$ OR CHARACTERISTIC$ OR MANUFACTUR$$$ OR PRODUC$$$$

CON1 A3 AND A4

Explanation

"ADJ" in A1 requires that both "SYNTHETIC" and "FOAM$" be close together in the order shown, to produce a hit. A3 indicates that either A1 or A2 are acceptable.

CON1 states that A3 and A4 may occur at any place in the same record to meet the information need. Only one type of logical connector is used in any one concept.

(3) Narrative statement. The same as under (1) and (2).

Profile

A1 SYNTHETIC OR PLASTIC ADJ FOAM$
A2 PROPERT$$$ OR CHARACTERISTIC$ OR MANUFACTUR$$$ OR PRODUC$$$$
CON1 A1 AND A2

Explanation

This is the concise way of setting up a profile from the statement given.

"OR" logical connector may be used with "ADJ" or "WITH" in the way shown in A1. (In the search expression CON1 you may use only logical connector "AND" between A1 and A2. "WITH" and "ADJ" could be used if A1 and A2 contained words connected by "OR.")

(4) Narrative statement. The same as under (1), (2), (3) but we do not wish to receive the information as far as marketing is concerned (and some other related terms).

Profile

A1 SYNTHETIC OR PLASTIC ADJ FOAM$
A2 PROPERT$$$ OR CHARACTERISTIC$ OR MANUFACTUR$$$ OR PRODUC$$$$
CON1 A1 AND A2
CON2 NOT MARKET$$$ OR SALE$ OR BUY$$$ OR CONSUM$*

Explanation

CON2 contains "NOT" which excludes all documents dealing with MARKET$$$ as well as other terms specified. These documents will not be matched by the profile.

(5) Narrative statement. The same as above; however, we request any information regarding polyurethane(s).

Profile

A1 SYNTHETIC OR PLASTIC ADJ FOAM$
A2 PROPERT$$$ OR CHARACTERISTIC$ OR MANUFACTUR$$$ OR PRODUC$$$$
CON1 A1 AND A2
CON2 NOT MARKET$$$
CON3 ABS POLYURETHANE$

Explanation

CON3 contains "ABS" logic. This means that any document dealing with "POLYURETHANE($)" will be picked out for the user. It overrides any other logic used.

6. CONCLUSION

The constructing of profiles in the TEXT-PAC system is more involved than in some of the less sophisticated IR systems, but it is more rewarding. Full text searching allows us to improve performance. We can obtain either better relevance or better recall, whatever our users prefer.

It should also be noted that questions for retrospective search are formulated in the same way as profiles in current awareness search.

NSL Profiling and Search Editing

by Georg Mauerhoff, Library
University of Saskatchewan

Introduction

As soon as the National Science Library made its Selective Dissemination of Information service commercially available in April, 1969, the Saskatoon campus of the University of Saskatchewan became involved in the CAN/SDI Project as a regional processing center. It is cumbersome for seekers of information to easily purchase individualized information services from remote search centers. Moreover, it is generally time-consuming and more demanding to participate in such a system. For expediency, a sufficiently large number of requests from scientific and technical personnel are dealt with locally and then searched remotely. Because of these factors, the idea of a regional processing center was initiated. This type of situation, a very common one now, permits the best kind of interaction between a large information system such as NSL's and its users.

A great many requests are still sent to Ottawa by mail or even phoned in, with even a few requestors visiting NSL and negotiating searches. On the whole though, interaction between system and user is achieved by a librarian or search editor who is located remote from the actual search center. As of about six months ago, over 80 subscriber communities availed themselves of the search service in this way. In other words, over 80 user groups across Canada rely upon a member of their own staff to perform profiling and search editing for them.

Profiling and Search editing are established professional operations, which are of major importance to both the operating efficiency and the economic efficiency of any dissemination system. In a system such as the CAN/SDI Project (see Brown, 1) these tasks play a more vital role, owing to the fact that they are decentralized and voluntary. A network of over 170 search editors are active across the nation, handling the output components of the system such as search statements, question analyses, and search strategies. Consequently, the part played by the editors must be looked upon by management as "the overwhelming variable, the major influencing factor" (2) affecting the performance of the system. It follows then, that anyone becoming a subscriber group must emphasize the need for locally establishing a position of profiler/search editor.

The profiler/search editor represents the system and is required to perform the interface with potential subscribers. He must therefore be familiar with the four document data bases which precipitate searches of the users' profiles, he must be aware of the indexing languages employed, and he should possess skills in the use of the various characteristics of the tape service used. Most importantly, a rigorous program of public relations should be waged by him in order to allow those with a recurring information need and those with an as yet undiscovered need, to make use of the service.

The Subscriber and his Profile

The many subscribers served by the CAN/SDI Project comprise various categories: individual scientists and engineers, each with his own peculiar information need(s); groups of individuals with overlapping interests; and organizations such as hospitals, government departments, and industrial firms. All, however, comprise a broad spectrum of subject competence - computer programmers and biochemists, pathologists and physicists, biologists and engineers, to name only a few. Each category of subscriber, it must be understood, also depicts a totally different kind of request in terms of depth of coverage and expansiveness.

The process of developing an interest profile for computer searching begins with a discussion of the CAN/SDI Project itself, and how the subscriber can derive benefits from the service. Referred to as user education in figure 1, the task requires participation by both the user and the search editor, and can take anywhere from 30 minutes (.50 hr.) to an average of 1.75 hour of their time. At this time, the user is shown sample profiles, typical printouts, (see Fig. 3 & 7) and is given an explanation of how searches can be conducted. Following system initiation, the subscriber is asked to submit a written description of his area of interest as it is related to his research projects, teaching or interests. The narrative (see Fig. 4) should employ language from the recent literature, and should contain a list of references using words which are to appear in the interest profile. The subscriber is also encouraged to draft a "raw" profile using the various search keys available on the data bases. These may be such things as authors, organizations, journals, words, cited questions, etc. Such preparation can require about 1 to 3 hours of the user's time.

The next task, one which again requires both user and search editor to be present, is referred to as "raw" profile analysis. Here, the person's interests are discussed and carefully analyzed, checks being made of the search terms, their frequency of usage, and their spelling. Finally, during the verification process, any additional word concepts which may have been brought about in the course of checking are incorporated into or deleted from, the documented profile. Probably the most important step in the profile preparation schedule, this part of the overall interface is also the most time-consuming. It requires at least 30 minutes (.50 hr.) from the user and 66 min. (1.10 hr.) from the editor. On the average, the times are usually 1.80 hr. and 2.20 hrs. respectively, really not that much when one considers this to be a one-time operation, i.e. barring any drastic changes in subjects.

Following this, the "final" profile design is created, coded, and submitted to NSL for processing (see Fig. 5). A turn-around time of about two weeks is customary for the first search results. Anywhere from two to an average of four updates will also be required over the subscription year in order to accommodate changing interests, new terminology and any problems with the profile which may have been discovered during the weekly and monthly statistical analyses (see Fig. 2).

Details of the CAN/SDI Search Programs and their Relationship to COMPENDEX, Chem. Titles and Retrospective

For processing reasons and for expediting the construction of a profile, five search keys or search units were developed by NRC's Computation Centre (see Fig. 8). Since the data bases are all translated into a MARC-like format, these search keys, where available, are equally applicable, whether the tape is from the Institute for Scientific Information (ISI) or from the Chemical Abstracts Service (CAS). The keys are personal author, corporate author, Coden, title and keyword, with coordination possible between or among any of these keys.

The title and keyword search keys consist of two types, and are permitted to have a term length of 40 characters, even though the average condensates search term length is only 9.6 characters (see Schwartz, 3). The types are:

1) Single words, such as OXYGEN and BRAIN.

2) Phrases, such as BLOOD BRAIN BARRIER and INHIBITORY TRANSMITTER.

The remainder of the keys require varying term lengths. These are given in the profile manual, e.g. a personal author on the ISI tape necessitates one knowing only up to 8 characters of the surname. Initials are optional.

The logic operators (see also Fig. 8) that join these five keys together are:

1) OR logic (/): a search will produce a reference containing <u>any</u> one or all members of a group of terms.

```
       BRAIN
    OR CENTRAL NERVOUS SYSTEM
    OR CEREBELL
    OR CEREBRAL
    OR SYNAP
    .
    .
    .
```

2) AND logic (&): a search <u>must</u> hit on a particular combination of terms to produce a reference.

```
                                 BRAIN
       SUBCELLULAR     AND OR    CENTRAL NERVOUS SYSTEM
    OR SUB-CELLULAR        OR    CEREBELL
                           OR    CEREBRAL
                           OR    SYNAP
                           .
                           .
                           .
```

3) NOT logic (¬): a search will exclude references if containing a particular term(s) as specified.

```
       OXYGEN      AND   HYPERBAR   ¬PATENT
    OR O2
```

4) THROUGH logic (—>): used to compress a grouping of single

words or phrases, and simplify the coding procedures.

OR RNA
. DNA
.
. ↓
.
.
OR CHEMISTRY

Complementing the logic operators is an array of quite sophisticated tools.

1. Masking or truncating of single words or phrases. This can take four forms: (A) LEFT truncation (*WORD) will result in a search for all words with the same root regardless of prefix.
   (B) RIGHT truncation (WORD*) will produce all words with the same root regardless of suffix.
   (C) LEFT-RIGHT truncation (*WORD*) will produce all words with the same root regardless of prefix or suffix.
   (D) NO truncation (WORD) will match all words only with the same root.
2. The Use of Weights

   If the AND/OR/NOT/THROUGH logic is incapable of providing the right degree of specificity, weighting procedures can be used.

For instance, in a 2-parameter search, it may be necessary to rule out certain combinations of terms and prevent their printing out, e.g. in the search $(A|B \longrightarrow G)\&(H|I \longrightarrow M)$, it may be necessary to prevent (D&H) and (E&K) from producing a hit. The trick then is to set a threshold weight of 10, and to give A, B, C, F, G, I, J, L, and M weights of 9, and D, E, H, and K weights of 1. All combinations except D&H and E&K will produce a search weight of 10 or more, and be printed out. The other two would always have a weight under the threshold.

You will notice that I have tried to relate NSL's tape service search characteristics wherever possible to the COMPENDEX, Chem. Titles and Retrospective systems.

The Search Editor's Tasks

Because of the various tasks involved in building a profile and due to the outlay in time, search editors should try to ascertain what kind of a work load they can maintain. In so doing, they will in the long run be constructing well-defined profiles and at the same time assuring the user of good returns. Each profile will be attached from the same point of reference, and enjoy the editor's same careful scrutiny.

According to figure 1, the minimum time required by an editor to prepare a profile for processing is 1.75 hours. On the average, however, 3.80 hours will be devoted to each user, with NSL throwing in an additional .5 or .6 hours for each and every profile. Overall search editing times are more revealing. Figures 1 and 2 show that over a period of one subscription year, a search editor requires on the average 25.40* hours to achieve a satisfactory profile. He is therefore able to provide a good service for as many as 35 profiles per year if he allows himself 235 working days at 3.75 hours a day. If, however, profiles are simpler and require less modification, the search editor could undertake a work load consisting of at least 108 profiles. This is a variable, though, and dependent on the nature of the profiles.

Assuming that you have the time, the user's narrative, and his "raw" profile, the task of analyzing and designing the final version is the only major task left. It is vital that this difficult process be treated properly and be recognized as "costly, time-consuming, elaborate, tedious, error-prone" (2), since the editor has to determine realistically the manner in which the profile is to be used. The searches must be defined in every detail, and must relate to the various concept levels. After all, the quality of the output is directly proportional to the quality of the profile.

Once the basic structure has been determined, i.e. how the concepts are to be linked, if at all, expansion of the concepts is undertaken in order to cover all the terms which may be found in

---

*= prep'n. & tab'n. of stats. (8 mo.) & analy. of stats. (8 mo.) & revisions (4x)
= 3.80 hr. & (32 wk. x .15 hr./wk.) & (8 mo. x 1.60 hr./mo.) & (4 rev. x 1 hr./rev.)
= 3.80 hr. & 4.80 hr. & 12.80 hr. & 4.00 hr.
= 25.40 hr.

relevant titles (see Fig. 6). Question formulation follows, and efficient search parameters are obtained. It must also be kept in mind that the subscriber determines the final form of the profile.

Data Bases

The NSL data bases presently consist of tape packages from three sources, which are generally discipline-oriented (see Mauerhoff, 4). Chemical Abstracts Condensates and Chem Titles both deal with chemistry; INSPEC deals with computers and control, physics, and electrotechnology. The Institute for Scientific Information's tapes, on the other hand, are embracing all disciplines in science and technology and must be attached in a slightly different manner. Thus, a point to remember is - get to know your data bases, their content and language, because they will more than likely aid in the construction and handling of interest profiles.

Take for example accession numbers (AN). These appear in the lower left hand corner of the printouts and coincide with printed indexes. On the Chem, Condensates tapes, the AN is the abstract number found in the weekly printed version of CA, with the same volume and number as the tape. On INSPEC tapes, the same is true. On the ISI tapes, however, the AN is the Original Article Tearsheet Service Number, which allows one to order a hard copy from ISI.

The Economics of Search Editing

The various tasks required for profiling represent an investment of anywhere from $2500 to $3750 per year in search editor salary time for the 35 to 108 profiles. The former is based on an annual salary of $5000, with the search editor working on SDI an average of 3.75 hours per day for 235 working days, approximately half-time. The latter assumes a salary of $7500 per annum (p.a.).

A technical man's salary can be thought of as ranging between $10,000 and $15,000 per annum. Thus "if one considers that the SDI service is capable of saving 1 per cent of the technical man's time" (5) or cut down his searching of the literature by 5 per cent (see Mohler, 5), significant savings will be realized.

Considering that a profile usually serves two or more researchers or technicians, one regional search editor who handles 35 to 108 profiles can be regarded as doing a very important job. He will be looking after the information needs of anywhere from 70 to 216 technicians. Their salaries can amount to $700,000 (70 men x $10,000 p.a.) or even as much as $3,240,000 (216 x $15,000 p.a.).

Of those salaries, anywhere from 1 per cent to 5 per cent will be spent each year on acquiring bibliographical information. This amounts to $7,000 per year, and can even go as high as $165,000 for those using all 5 per cent.

Were these same researchers to utilize an SDI service such as NSL's, or AIRA's, or COMPENDEX, subscriptions would cost them $3500 for 35 profiles at $100 per profile. At most, subscriptions could cost $14,040 for 108 profiles at $130 per profile. There would also be some small expenditure in time for the users, but this can be regarded as negligible.

Overall savings to the users can be computed by taking the difference between investment costs, i.e. search editor time plus profile subscriptions, and annual information acquiring costs. The difference in costs could be as little as $1000, or in the maximum case $144,210. The savings, it must be remembered, are not cash savings, but merely a displacement of time, since the time saved will be reallocated. In addition, more timely information is being brought to the user, which also could reflect savings to management.

Conclusion

This description of profiling and search editing may seem complex at first, but once one delves into the manual and initiates several profiles, one can soon become quite facile with even the most complicated aspects.

NSL Profiling and Search Editing
Bibliography

1. Brown, J. The CAN/SDI Project: The SDI program of Canada's National Science Library, Special Libraries (October 1969) 501-509.

2. Saracevic, T. An inquiry into testing of information retrieval systems: analysis of results. Center for Documentation and Communication Research, School of Library Science, Case Western Reserve University, Cleveland, Ohio, 1968, 204 p. (Comparative Systems Lab. Technical report, part II)

3. Schwartz, E.S.; Williams, M.E.; Schipma, P.B.; Mankus, B.L. Educational and commercial utilization of a chemical information center. Quarterly progress report no.6. IIT Research Inst., Chicago, Ill., 25 Dec. 1969, 46 p. (PB-191 550)

4. Mauerhoff, G.R. The CAN/SDI Project: User Populations and Data Bases. Journal of Information Science in Canada, Summer (July 1970) in press.

5. Hoshovsky, A.G.; Downie, C.S. Selective dissemination of information in practice: survey of operational and experimental SDI systems. Presented at the 1967 Congress of the International Federation on Documentation (FID), Tokyo, September, 1967.

6. Mohler, I.C. The impact of information science on biology; a possible society role. Journal of the Washington Academy of Science, 59 (September 1969) 117-119.

Prepared by
G. R. Mauerhoff
Search Editor, CAN/SDI
Office of Technical Services
Library
University of Saskatchewan
Saskatoon, Canada.

NSL PROFILING AND SEARCH EDITING AT THE LIBRARY
UNIVERSITY OF SASKATCHEWAN, SASKATOON

Fig. 1. **Interest profile preparation schedule**

Task	Hours of Elapsed Time					
	Minimum			Average		
	User	Editor	NSL	User	Editor	NSL
1. User education	.25	.25		.75	1.00	
2. "Raw" profile design	1.00			3.00		
3. "Raw" profile analysis						
(a) Discussion	.25	.25		.50	.50	
(b) Word frequency check		.10			.10	
(c) Word usage study		.40			.80	
(d) Word spelling check		.10			.10	
(e) Verification	.25	.25		1.30	.70	
4. "Final" profile design						
(a) Question formulation		.10			.30	
(b) Coding		.10			.10	
(c) Typing		.20			.20	
5. Profile processing						
(a) Verification			.30			.40
(b) Keypunching			.10			.10
(c) Testing			.10			.10
(d) Running						

NSL PROFILING AND SEARCH EDITING AT THE LIBRARY, UNIVERSITY OF SASKATCHEWAN, SASKATOON

Fig. 2. Interest profile updating schedule

Task	Hours of Elapsed Time					
	Minimum			Average		
	User	Editor	NSL	User	Editor	NSL
1. Printout evaluation •						
(a) relevance judgments	.10			.25		
(b) analysis of references	.10			.15		
2. Tabulation of feedback statistics •						
(a) by equation		.02			.05	
(b) by search term(s)		.05			.10	
3. Analysis of statistics & user's notes ˇ						
(a) isolating inefficient search term(s)		.20			.80	
(b) reformatting questions; addition(s), deletion(s), change(s), etc.		.20			.80	
4. "Final" profile design						
(a) verification		.30			.60	
(b) question formulation		.05			.20	
(c) coding		.05			.10	
(d) typing		.10			.10	
5. Profile processing						
(a) verification			.20			.20
(b) keypunching			.05			.10
(c) testing			.05			.10

• weekly

ˇ monthly

Fig. 3. CAN/SDI Output as used at the Library, University of Saskatchewan, Saskatoon.

```
BLOOD BRAIN

JEPPSSON P G UNIV HOSP LUND,DEPT NEUROL, LUND, SWE
OLIN T

NEUROTOXICITY OF ROENTGEN CONTRAST MEDIA - STUDY OF BLOOD-BRAI
N BARRIER IN RABBIT FOLLOWING SELECTIVE INJECTION OF CONTRAST
MEDIA INTO INTERNAL CAROTID ARTERY

ACTA RADIOLOGICA. DIAGNOSIS
VOLUME 10, ISSUE  N1, YEAR 1970, PAGE 17, REF 035


IS THIS CITATION USEFUL?   YES   [✓ NO]   CANNOT TELL    COMMENT
AN 0007625 P 0886 EN 01 TW 000 WT 000 S I1470 TP ARTC L ENG
```

Title

```
HYPOXI, BRAIN

BEKETOV A I   KRYM. MED. INST., SIMFEROPOL, USSR).
SAREGIN D I

EFFECT OF HEPARIN ON BLOOD AERATION AND OXYGEN STRESS IN THE B
RAIN AND MUSCLES UNDER HYPOXIA CONDITIONS.

GEPA C*
VOLUME 1969, PAGE 42-4

OXYGEN ** BLOOD ** HEPARIN ** BRAIN ** O ** MUSCLE **

IS THIS CITATION USEFUL?  [✓ YES]   NO    CANNOT TELL    COMMENT
AN 1095680 P 0886 EN 09 TW 000 WT 000 S C2172 TP CONF L
```

Keyword Abstract

```
C88, LAB, CLINIC

LOTT, J.A.(OHIO STATE UNIV. HOSPITAL, COLUMBUS, USA)
GRUMER, H.-D.; SCOTT, J.A.;

EARLY EXPERIENCE WITH THE LINC-8 IN A HIGH VOLUME CLINICAL CHE
MISTRY LABORATORY

DECUS PROCEEDINGS OF THE SPRING SYMPOSIUM 1969
PP: 301-2 , 1969

THE EXPERIENCE WITH THE LINC-8 USING THE WISCONSIN PROGRAM IN
A HIGH VOLUME CLINICAL CHEMISTRY LABORATORY IS DISCUSSED. THE
USE OF THE CONVERSATIONAL MODE BY LABORATORY PERSONNEL, ON-LIN
E AND OFF-LINE CALCULATIONS, MODIFICATIONS OF AUTOANALYZERS FO
R IMPROVED PERFORMANCE WITH THE LINC, AND ACCEPTANCE OF THE CO
MPUTER IN THE LABORATORY ARE INCLUDED


WAKEFIELD, MA. USA , 12-13 MAY 1969

IS THIS CITATION USEFUL?  [✓ YES]    NO    CANNOT TELL    COMMENT
AN 09666    P 0946 EN 08 TW 000 WT 000 S P1270 TP PROC L ENG
```

Abstract

Fig. 4. Coding sheet with narrative and references.

PROFILE - NUMBER

886

INSERT YOUR ADDRESS LABEL IN THIS BLOCK

J. D. Wood,
Department of Biochemistry,
University of Saskatchewan,
Saskatoon, Saskatchewan.

SHEET- NUMBER

1

STATE YOUR SEARCH REQUEST IN NARRATIVE FORM. ADD TWO REFERENCES OF PAPERS PUBLISHED BY YOU OR A COLLEAGUE WORKING IN YOUR FIELD. (PLEASE TYPE OR PRINT)

I am interested in the field of neurochemistry, particularly in the biochemical mechanisms involved in the production of convulsions.

References:

(1) Jamieson, D. and Van Den Brenk, H.A.S.: The Effects of Antioxidants on High Pressure Oxygen Toxicity. Biochemical Pharmacology, 1964, vol. 13, pp.159-164.

(2) Van Den Brenk, H.A.S. and Jamieson, D.: Brain Damage and Paralysis in Animals Exposed to High Pressure Oxygen - Pharmacological and Biochemical Observations. Biochemical Pharmacology, 1964, vol. 13, pp. 165-182.

(3) Bean, J.W.: Cerebral O_2 in Exposure to O_2 at Atmospheric and Higher Pressure, and Influence of CO_2. Reprinted from American Journal of Physiology, v.201, no.6, December 1961.

(4) Shilling, C.W. and Adams, B.H.: A Study of the Convulsive Seizures Caused by Breathing Oxygen at High Pressures. U.S. Naval Medical Bulletin, v.31, 1933.

(5) Graham, L.T., Jr., Shank, R.P., Werman, R. and Aprison, M.H.: Distribution of Some Synaptic Transmitter Suspects in Cat Spinal Cord. Glutamic Acid, Aspartic Acid. -Aminobutyric Acid, Glycine and Glutamine. Journal of Neurochemistry, 1967, vol. 14, pp. 464-472.

(6) Wallach, D.P.: Studies on the GABA Pathway -I The Inhibition of γ-Aminobutyric Acid -α- Ketoglutaric Acid Transaminase In Vitro and In Vivo by U-7524 (Amino-Oxyacetic Acid). Reprinted from Biochemical Pharmacology, vol. 5, no.4, pp. 323-331.

LIST PROFILE WORDS AND SEARCH EXPRESSIONS ON REVERSE SIDE

Fig. 5. "Final" profile design

TT	W	AC	PROFILE WORDS	TT	W	AC	PROFILE WORDS
A		A	KRNJEVIC K	T		AM	O2
A		B	ROBERTS E	T		AN	HIGH *
T		C	*AMINOBUTYR*	T		AO	PRESSURE*
T		D	GABA	T		AP	HYPERBAR*
T		E	HYDRAZIDE *	T		AQ	*CONVULS*
T		F	HYPEROXIA	T		AR	SEIZURE*
T		G	INHIBITORY TRANSMITTER	T		AS	SUB-CELLULAR
T		H	ISONIAZID*	K		AT	PATENT
T		I	SEMICARBAZIDE *	T		AV	MATUR *
T		J	THIOCARBOHYDRAZIDE*	T		AW	PYRIDOX*
T		K	THIOSEMICARBAZIDE*	T		AX	PICROTOXI*
T		L	BLOOD BRAIN	T		AY	STRYCHNIN*
T		M	LEARNING	T		AZ	ACID*
T		N	MEMORY				
T		O	*RNA				
T		P	DNA				
T		Q	*NUCLEIC ACID*				
T		R	PROTEIN*				
T		S	*CHEMISTRY				
T		T	EXCIT*				
T		U	HYPOXI*				
T		V	INHIBIT*				
T		W	SUBCELLULAR				
T		X	GLUTAM*				
T		Y	*GLUTAR*				
T		Z	BRAIN*				
T		AA	CENTRAL NERVOUS SYSTEM*				
T		AB	CEREBELL*				
T		AC	CEREBRAL				
T		AD	*SYNAP*				
T		AE	AMINO*				
T		AF	AMINOACID*				
T		AG	RELEAS*				
T		AH	TRANSPORT				
T		AI	UPTAKE				
T		AJ	DECARBOXYLASE*				
T		AK	TRANSAMINASE				
T		AL	OXYGEN*				

EN	S	NR	TW	SEARCH EXPRESSIONS
1	I	99		A|B –> D|F|G|L
2	I	99		(M|N)&(O|P –> S)
3	I	99		(T|U –> Y|AS)&(Z|AA –> AD)
4	I	99		(X|AE&AZ|AF)&(E|H –> K|(AG|AH|AI)&(Z|AA –> AD))
5	I	99		AJ&(X|Y)|Y&AK|(AL|AM)&(AN&AO|AP)
6	I	99		(AQ|AR)&(E|H –> K|X|AE&AZ|AF|AV –> AY)
7	C	99		(A|B –> D|F|G|L)¬AT
8	C	99		(M|N)&(O|P –> S)¬AT
9	C	99		(T|U –> Y|AS)&(Z AA –> AD)¬AT
10	C	99		(X|AE&AZ|AF)&(E|H –> K (AG|AH|AT)&(Z|AA –> AD))¬AT
11	C	99		(AJ&(X|Y)|Y&AK|(AL|AM)&(AN&AO|AP))¬AT
12	C	99		(AQ|AR)&(E|H –> K|X|AE&AZ|AF|AV –> AY)¬AT

Fig. 6. Documenting search requests

```
1.  A  KRNJEVIC K                                    6.[ AQ *CONVULS* ]   &  [ AE  AMINO* & AZ  ACID*
    B  ROBERTS E                                       [ AR  SEIZURE* ]      [ AF  AMINOACID*
    C *AMINOBUTYR*                                                           [  X  GLUTAM*
    D  GABA                                                                  [  E  HYDRAZIDE*
    F  HYPEROXIA                              ¬ AT                           [  H  ISONIAZID*
    G  INHIBITORY TRANSMITTER                                                [  I  SEMICARBAZIDE*
    L  BLOOD BRAIN [BARRIER]                                                 [  J  THIOCARBOHYDRAZIDE*
                                                                             [  K  THIOSEMICARBAZIDE*
                                                                             [ AV  MATUR*
                                                                             [ AW  PYRIDOX*
                                                                             [ AX  PICROTOXI*
2.[ M  LEARNING ]   &   [ S *RNA              ]                              [ AY  STRYCHNIN*  ¬ PATENT ]
  [ N  MEMORY   ]       [ P  DNA              ]
                        [ Q *NUCLEIC ACID*    ]   ¬ PATENT
                        [ R  PROTEIN*         ]
                        [ S *CHEMISTRY        ]

3.[ T  EXCIT *       ]         [ Z  BRAIN*                  ]
  [ U  HYPOXI*       ]         [ AA CENTRAL NERVOUS SYSTEM* ]   ¬ PATENT
  [ V  INHIBIT*      ]    &    [ AB CEREBELL *              ]
  [ W  SUBCELLULAR   ]         [ AC CEREBRAL                ]
  [ AS SUB-CELLULAR  ]         [ AD *SYNAP*                 ]
  [ X  GLUTAM*       ]
  [ Y *GLUTAR*       ]

4.[ AE AMINO* AZ ACID * ]  &  [ E  HYDRAZIDE*                                              ]
  [ AF AMINOACID *      ]     [ H  ISONIAZID*                                              ]
  [ X  GLUTAM *         ]     [ I  SEMICARBAZIDE *                                         ]
                              [ J  THIOCARBOHYDRAZIDE *                                    ]   ¬ PATENT
                              [ K  THIOSEMICARBAZIDE *                                     ]
                              [                                                            ]
                              [ [ AG RELEAS *   ]   &  [ Z  BRAIN *                    ]   ]
                              [ [ AH TRANSPORT  ]      [ AA CENTRAL NERVOUS SYSTEM *   ]   ]
                              [ [ AI UPTAKE     ]      [ AB CEREBELL *                 ]   ]
                              [                        [ AC CEREBRAL                   ]   ]
                              [                        [ AD *SYNAP*                    ]   ]

5.  AJ  DECARBOXYLASE *  &  [ X  GLUTAM* ]
                            [ Y  GLUTAR* ]   ¬ PATENT

    X *GLUTAR*  &  AK  TRANSAMINASE

  [ AL OXYGEN* ]  &  [ AN HIGH*  &  AO PRESSURE * ]
  [ AM O2      ]     [ AP HYPERBAR*               ]
```

Fig. 7. Computer version of profile

```
                 DATE: MAY 14, 1970
       U 886                 020270 PROFILE  WOOD J D
RECORD NOT ON FILE
        A          A KRNJEVIC K
        A          B ROBERTS E
        T          C *AMINOBUTYR*
        T          D GABA
        T          E HYDRAZIDE*
        T          F HYPEROXIA
        T          G INHIBITORY TRANSMITTER
        T          H ISONIAZID*
        T          I SEMICARBAZIDE*
        T          J THIOCARBOHYDRAZIDE*
        T          K THIOSEMICARBAZIDE*
        T          L BLOOD BRAIN
        T          M LEARNING
        T          N MEMORY
        T          O *RNA
        T          P DNA
        T          Q *NUCLEIC ACID*
        T          R PROTEIN*
        T          S *CHEMISTRY
        T          T EXCIT*
        T          U HYPOXI*
        T          V INHIBIT*
        T          W SUBCELLULAR
        T          X GLUTAM*
        T          Y *GLUTAR*
        T          Z BRAIN*
        T         AA CENTRAL NERVOUS SYSTEM*
        T         AB CEREBELL*
        T         AC CEREBRAL
        T         AD *SYNAP*
        T         AE AMINO*
        T         AF AMINOACID*
        T         AG RELEAS*
        T         AH TRANSPORT
        T         AI UPTAKE
        T         AJ DECARBOXYLASE*
        T         AK TRANSAMINASE
        T         AL OXYGEN*
        T         AM O2
        T         AN HIGH*
        T         AO PRESSURE*
        T         AP HYPERBAR*
        T         AQ *CONVULS*
        T         AR SEIZURE*
        T         AS SUB-CELLULAR
        K         AT PATENT
        T         AV MATUR*
        T         AW PYRIDOX*
        T         AX PICROTOXI*
        T         AY STRYCHNIN*
        T         AZ ACID*
       E01 I 99      A|B-D|F|G|L
       E02 I 99      (M|N)&(O|P-S)
       E03 I 99      (T|U-Y|AS)&(Z|AA-AD)
       E04 I 99      (X|AE&AZ|AF)&(E|H-K|(AC|AH|AI)&(Z|AA-AD))
       E05 I 99      AJ&(X|Y)|Y&AK|(AL|AM)&(AN&AO|AP)
       E06 I 99      (AQ|AR)&(E|H-K|X|AE&AZ|AF|AV-AY)
```

Fig. 8. Summary of Tape Service Search Characteristics

Characteristic	Tape Services					
	National Science Library (NSL)				U. of Calgary	AIRA
	ISI	CAC	Chem. Titles	INSPEC	COMPENDEX	Chem. Titles & Retrospective
1. Type of Searching						
(a) Current awareness - CA	X	X	X	X	X	X
(b) Retrospective - R					X	X
2. Frequency of Search	Weekly	Weekly	Bi-weekly	Bi-weekly	CA - Monthly R - ○	CA - Bi-weekly R - On demand
3. Coverage (records)	5000-15000	5000	5000	4000	CA - 6000 R - Jan.'69-	CA - 5000 R - Jan.'63-
4. Searchable units						
(a) Personal Author - PA	X	X	X	X	X	X
(b) Corporate Author	X	X		X	X	
(c) Citations	X					
(d) Subject heading/code				X	X	
(e) Source/Coden	X	X	X	○	X	X
(f) Title	X	X	X	X	X	X
(g) Keyword/access Word/						
(h) Index Line		X		○	X	
(i) Abstract					X	
5. Search Techniques						
(a) OR logic	X	X	X	X	X	X
(b) AND	X	X	X	X	X	X
(c) WITH					X	
(d) ADJ	X	X	X	X	X	
(e) NOT	X	X	X	X	X	X
(f) ABSOLUTE					X	
(g) CONTROL	Equivalent to the Term Type (TT)				X	
(h) NOT-CONTROL					X	
(i) THROUGH	X	X	X	X		
(j) Term & Threshold Weights	X	X	X	X		X
6. Extended Characteristics						
(a) Capitalization	Upper Case (UC)	UC	UC	UC	UC & LC	UC
(b) Truncation (Left - L, Right - R)	L & R Uncond'l.	L & R Uncond'l. except keywords	L & R Uncond'l.	L & R Uncond'l. except subject codes	Right - Select 6 characters R Uncond'l.	L & R Uncond'l.
(c) ¬ PATENTS		X				
(d) Types of Logical Connectors in concept/expression	3	3	3	3	1	3
(e) Length profile term & length of search term	40 (40) char. except PA	40 (40) char. except PA	40 (40) char. except PA	40 char. except PA	38 (20) char.	
(f) Back Referencing	Equivalent to using THROUGH logic				X	
(g) Output option	X	X	X	X		
(h) Tape Search Option	X	X	X	X		
(i) "2 up" print option	X	X	X	X		
(j) Feedback option	X	X	X	X		
(k) Output option, i.e. paper, card or tape	X	X	X	X		

○ will be searchable at a later date.

Table 1. (Estimated) Annual Investment Costs Per Region

	Minimum	Maximum
Search editor (half-time for 235 days) at $5,000 per year $7,500	$2,500	$3,750
Profile Subscriptions (35 to 108)	$3,500	$14,040
Total$_1$	$6,000	$17,790

Table 2. (Estimated) Annual Information Acquiring Costs Per Region

Technical man (1% of $10,000 salary x 70 users)	$7,000	
(5% of $15,000 salary x 216 users)		$162,000
Total$_2$	$7,000	$162,000
Savings (Total$_2$ - Total$_1$)	$1,000	$144,210

Discussion of Experiences of Those Providing, and Those Using, SDI Services

K.E. Marshall
Chairman

Discussion Of The Experiences Of Those Providing, And Those Using SDI Services

K.E. Marshall, Chairman

MARY FLETCHER: Alberta Information Retrieval Association

The following services are available to us at AIRA in answering an inquiry:

Service	Basic Applications
* Compendex	Industry, Research
* Chemical Titles	Science, Education
* Chemical Titles Retrospective Searches	Science, Research
* ISI (via NSL)	Industry, Science
* CA Condensates (via NSL)	Science, Research
* INSPEC (via NSL)	Industry, Science
TIS (via RCA, NRC)	Industry
TD (via NRC)	Industry
* KWIC (via I.E.S.)	Industry
ICURR	Government
TAR SANDS	Government, Industry

(*indicates a computerized service)

The first three services listed and the last two are available directly. The Association acts as a middleman in the use of the other services by passing the inquiry to the organization indicated in parentheses.

Without doubt, the main problem as providers of SDI is the old communication problem. Profiles are often submitted by users in such a form as to be unsuitable for use on a computerized data base. This problem can only be resolved by extensive education of all persons using SDI services.

We do have some minor problems, as does any project of this type in its initial stages of development. These problems are all being looked into and every possible effort is being made to resolve them.

We offer to any user of the Compendex service a three-month, no obligation trial period. If, after this time, the subscriber is satisfied with the service, he will be billed from the start of the trial period.

It is hoped that in the future the scope of the services offered will be broadened and that new areas of interest will be covered by new information retrieval services.

Abbreviations:

CA	Chemical Abstracts
ICURR	Intergovernmental Committee on Urban and Regional Research
I.E.S.	Industrial Engineering Services
INSPEC	Information Services in Physics, Electrotechnology and Control
ISI	Institute for Scientific Information
KWIC	Keyword in Context (Index)
RCA	Research Council of Alberta
NRC	National Research Council
NSL	National Science Library
TD	Technological Developments
TIS	Technical Information Services

BEVERLY CHANDLER: General Sciences Library, University of Alberta

The reference librarians of the General Sciences Library of the University of Alberta provide search editing for the SDI services offered by the National Science Library (CAN/SDI) and the University of Calgary (Compendex). We have attended workshop courses sponsored by NSL and the Alberta Information Retrieval Association (on Compendex). We have now compiled a dozen profiles with varying degrees of success.

The search editor is merely a middleman and must interface successfully with both the subscriber and the agency running the SDI service. There are several necessities required from the SDI agency:

(1) An Instruction Course:
Aside from the obvious learning value, such a course introduces the personalities behind the printout. This humanization allows freer interchange of queries. It also demonstrates to the search editor the potential and the pitfalls of the service.

(2) An Instruction Manual:
This, the back-up to the course, must be accurate and thorough, with provision for updating. It should supply some indication of the scope of each data bank, either by lists of subject headings or of journals currently scanned. Types of material (e.g. patents, government reports, etc.) covered should be listed. Updatings should mention new inclusions in the data bank. For our purposes sample printouts for each tape service, rather than mimeographed replicas, would be helpful.

(3) Professional Service:

When dealing with current awareness services prompt and courteous replies to queries are essential. Changes made in profiles by the service agency should be reported to both the subscriber and the search editor. These changes can help to educate the search editor.

The routine for a search editor/subscriber interaction should be:

(1) A preliminary conversation with the potential subscriber to explain SDI.

(2) The subscriber submits a narrative statement with sample references.

(3) The search editor translates this into SDI logic.

(4) Subscriber and search editor carefully examine the complete profile before mailing into the agency.

However, this routine will vary greatly depending on the subscriber's interest and/or faith in SDI and/or the search editor. Also, each subscriber has a unique use for SDI. In addition to current awareness or literature searching, other uses which we have encountered are: to update reading lists for an undergraduate engineering class; to select current material for a pollution-awareness library; and to maintain an ongoing bibliography.

Feedback should be a corollary to SDI, but it is not. However, since the subscriber must come to the General Sciences Library for the source document of his citation we do enjoy forced feedback.

Search editing is an extension of the regular information services of the reference librarians in the General Sciences Library. For us SDI is valuable because:

(1) It provides a painless and reasonably efficient access to the technical literature.

(2) It produces accurate bibliographic information for retrieval in our own library or through inter-library loan.

(3) It promotes contact between faculty and library staff.

(4) It forces evaluation of our collection.

Unfortunately, university faculty and students are generally unaware of the delights of SDI. We introduce the term "selective dissemination of information" wherever logically possible: to graduate students on library orientation tours; to faculty members submitting inter-library loan requests; and to anyone doing a comprehensive literature search.

We periodically blanket the science faculty with notices concerning current awareness and information retrieval. One prime target is the overworked graduate student, who could save many hours of work if he subscribed to SDI. He resists for many reasons, but chiefly because of the cost.

In spite of all our efforts, the best advertisement is by word of mouth, from a satisfied subscriber to his friends and colleagues. Now if we could only get that first satisfied graduate student ...

STEPHEN HOLLANDER: Science and Technology Librarian, University of Manitoba

At the University of Manitoba we rely largely on the National Science Library CAN/SDI Service and the Alberta Information Retrieval

Association's Compendex service. Since CAN/SDI was promoted at the University of Manitoba 18 months ago by Derek Francis and Boris Raymond, it has grown to become one of the largest such programs at any University in Canada. Because of the size of our SDI program we have a fairly good idea, I believe, of the types of problems which may occur.

One problem involves the subject approach. Some services scan the titles of articles and thus rely on descriptive titles to a large extent. Non-descriptive titles can, therefore, cause problems. For example, a paper published in the field of genetics was entitled ''Either Or', and there is no way that a profile based on title words could be designed in advance to retrieve it. The ambiguity of words creates problems also. One chemist is interested in atomic and molecular scattering collisions. Try as we may, we have not been able to suppress references to collisions of ships at sea, planes in the air, and cars on the road. This is the problem when one word has several meanings. On the other hand, you have problems when it takes many terms to express a single concept. This is especially true in the field of psychology, which has a largely uncontrolled vocabulary. In one profile for a psychologist, 22 terms were needed to express three concepts, and we are not certain that we have all of them yet.

Another problem arises in the area of phyletic classification. Biologists often express their interests in terms of large taxonomic groups, while papers generally deal with only one or two species. Unless the large group name is mentioned in the title there is no way in which the computer can place a genus in its appropriate phylum. For example, a researcher interested in Nematodes will say so in his profile. Unless every generic name is listed also, papers on Neoplactana or Panagrellus will not be retrieved. Fortunately, in such cases as these, one can usually pin the researcher down to the genera on which he is actually working, and while he would be pleased to have papers on, say, all rodents, he may be satisfied if he can have references to papers on rats and mice. This is sometimes true, but not always. In one profile there are nearly 100 terms for genera and species of algae. A similar problem may also arise in the field of chemistry. A researcher may be interested in heavy metals, but the computer does not know that tantallum, platinum, gold, etc., are heavy metals and so each must be listed. This together with the abbreviations and truncations necessary to allow for isotopes, makes the creation of a profile for this type of subject a complex job indeed. Truncations are a problem in themselves. Unpredictable describes the output from some truncations. One would think the EMBRYO* would be a garbage dump for an embryologist, but as it turns out on one profile that I have had, this is not the case. Truncation of STRAIN, both before and after with the intention of catching microstrain, microstrains and related terms also collects constrained, restraining, distrained, etc. The truncation GENE* opens up not only the field on genetics (which was wanted) but also brings in generators, general, etc.

A problem unique to the INSPEC tapes is the use of section and chapter codes in a search. If the search is restricted to certain sections, the search expression may be quite broad while still reflecting only the field of interest. However, using these headings evidently takes up a great many of the 1000 allowable operations in the search program. In one profiling 52 terms grouped into 9 search expressions fit quite nicely into the allowable 1000 operations. However, when the code A16* is added to the search expressions, it exceeds the 1000 operations. Thus the utility of the chapter and section headings is somewhat restricted.

Let us now get away from the problems and look at some of the potential. As it stands now the CAN/SDI Service is simply a citation retrieval operation. Expansion into more sophisticated functions would not be beyond the current state of the art. It would be possible, for instance, for the computer to note the statistically significant coincidence of certain words in their semantic and syntactic context with the profile words in relevant titles. Using this mechanism, the profiles could be adjusted by the computer to gain in relevance as time goes by. Also the report of these coincidences could be used in the preparation of an extremely sophisticated co-ordinate index which would show word clusters or, perhaps more appropriately, concept clusters as they occurred in relevant articles. This drawing together, or eliciting of relationships, would be a mechanization of the first intellectual step in the preparation of a research report or of a Thesaurus. It may go even farther, in that it could suggest lines of research that have possibly not been covered before. Thus it may be possible in this field to move from citation retrieval to genuine artificial intelligence.

GORDON THOMPSON: Syncrude Canada Ltd., Edmonton

There are a number of reasons for users to take SDI services. These are:

(1) A need to keep up-to-date.

SDI services give an early alert of papers of interest.

The user is relieved from the drudgery of hand searching CT, CA, Current Contents, Engineering Index or Science Abstracts volumes.

(2) Bibliographic Function.

The SDI output may be accumulated and filed according to subject to provide an up-to-date bibliography of the interest area.

(3) Time-saving Approach.

In industry especially, time is money. The cost of SDI services is more than returned in time-savings.

(4) The Vital Information Approach.

An important paper, especially in fast developing fields, may provide the key to large dollar savings. This requires thorough coverage of all available literature.

At Syncrude we have used CT, CA and Compendex services.

Chemical Titles services are provided by both NSL and AIRA. This service is very early with the printout of search results often arriving before the journal issues referenced. We have been taking CT search services from both NSL and AIRA, and we have found AIRA to be earlier than NSL in providing output.

I prefer the NSL format (cards) to the AIRA format. The cards are easy to file, and are referenced by the term(s) causing the hit which is printed at the top.

A major problem has been the difficulty of confirming our interest in a specific paper. We frequently find that papers with interesting titles are of little use. CA provides the abstract from which a judgment of the papers value may be made, but no similar backup material is available from CT hits short of obtaining a copy of the paper, which often proves very time-consuming.

We have had excellent results with the CT retrospective search service offered by AIRA. The question which we submitted resulted in little noise primarily because the terms are specific to our interest area. Other questions which we have searched in the current CT program would be hopeless to use in the retrospective search because of the large volume of output, most of which is relevant but with only a small percentage of hits of direct interest to us.

In the Chemical Abstracts Condensates search we had a program in which the same terms were asked for as title and keyword terms. We found the same percentage relevance from each but we had more hits with keywords than title words. We found very few useful hits picked up from the title search which were not picked up in the keyword search also. I like the output format from this service, especially the printing out of all keywords applied to the hit. This is handy in restricting a question to reduce noise. I have found some difficulty in choosing keywords and search expressions which will restrict without losing papers of interest. One suggestion by Miss Gaffney is that I use two search expressions for the same question, one designed to be well restricted followed by a very general expression to get all possible relevant papers. Since an item will be printed only once in each profile search the hits picked up by the restricted search are not repeated in the more general one. If useful references are found in the more general search the restricted search expression may be modified so that in future such items will be picked up. If nothing useful is picked up by the general search expression after a number of tapes have been searched, it may be dropped.

Before saying anything about our Compendex searches, I must admit that I am not familiar with the printed Engineering Index. At least part of my trouble may be due to this lack of familiarity.

I have found Compendex to be most frustrating. I feel it should be the best of the three services we use on the basis of providing and searching the abstracts, the coverage it gives to papers issued at meetings, and translated foreign journals. We started with an open search, then restricted our search to omit journal and author affiliation when those items proved to be sources of noise. We have continued to narrow the search but so far have not eliminated the noise. My major complaint at the moment is that the items are picked up on subject headings, etc., which are not printed out and there is no indication of what term caused the hit. Since hits are not printed out in an order related to the question this leaves the user guessing about the reason why an item was picked up. I feel that an indication of what caused the hit is a major requirement.

Finally a few over-all comments:

(1) There is a need for help in preparing profiles. This help needs to be close at hand so that an active exchange of ideas and material is possible.

(2) I feel there is a need for a different type of feedback.

(3) We need to define relevance so that everyone uses it with the same meaning.

(4) I think there is a place for an indication of whether the items are in the area of interest even if they are not useful.

General Discussion

Discussion followed the last of the formal presentations by the panel. Among the topics covered was a plea for some close co-operation in the format of the tapes produced by the various agencies. The preparation of search expressions for the different services would be simplified if the formats were similar. It was pointed out that most of the tape services which we use for SDI services were in the first instance produced to aid in the production of the indexes of the printed abstracting or indexing service. FRANK DOLAN gave the meeting a little background information on the steps which are being taken by a committee which has been formed to look into this very matter. It was also pointed out that some of the problems which currently face some of Mr. Hollander's biological search profiles may be partly, at least, solved when the CAN/SDI Service adds the BIOSIS (Biological Abstracts) tapes to its available data bases. (It is expected that this will become available early in 1971). The titles of papers are edited before being put on tape so that such larger groupings (phylum, order, etc). are included when the author has not already done so. This service, also additional indexing terms, are added to nondescriptive titles.

Keynote Address

Dr. Carlos A. Cuadra

On-Line Systems *

by Dr. Carlos A. Cuadra
Library and Documentation Systems Department
System Development Corporation
ASIS DISTINGUISHED LECTURER for 1970

Dr. Cuadra described his initial contact with on-line systems, in 1957, and then reviewed the major virtues and problems with on-line systems in library and information science. The potential virtues are speed, intimacy, and--if time-sharing is involved--economy. The major problems are the cost of the large-size computers and files necessary for bibliographic data, the high cost of communications, and the generally poor design of the user-system interfaces.

Dr. Cuadra discussed some of the key interface "provisions" in on-line retrieval systems and indicated how they are being used in a new system now being operated for the U.S. National Library of Medicine. He argued that, in addition to engineering the necessary capabilities into the system, system implementers must also try to engineer user acceptance. The most common pitfalls here include failure to take into account the social context of the user terminal and overselling the capabilities of the system.

Dr. Cuadra concluded his talk by posing several challenging issues for the U.S. (and Canadian) information science community, including the problem of deciding how the individual user is to learn how to cope with a diversity of on-line files and communication.

* This is only the abstract of Dr. Cuadra's talk. The full text will be published in the March-April issue of JASIS.

BUSINESS
MEETING

Election of Officers

The charter of the Western Canada Chapter of ASIS states: "There shall be a Nominating Committee, consisting of a chairman and a member from each participating province or territory designated by the Chapter Chairman. This committee shall present a list of nominees to the members of the Chapter, as hereinafter provided. The Nominating Committee, at its descretion, may present to the membership more than one candidate for any office."

H.S. Heaps as Chairman of the Western Canada Chapter designated in August, 1970, the following as members of the Nominating Committee.

David T. Wilder (Manitoba)
George Piternick (British Columbia)
Arlean E. McPherson (Saskatchewan)
G. Bert Reaburn (Yukon and Northwest Territories)
John Scott Truswell (Alberta)

The charter states that the Vice-Chairman shall automatically succeed to the office of Chairman and that the Secretary-Treasurer shall be elected for a period of two years. Thus Frank Dolan and Nita Cooke will automatically be Chairman, and Secretary-Treasurer, respectively for the period 1970-71.

This leaves the three offices of Vice-Chairman, chapter representative and alternate representative vacant.

Secretary-treasurer's Report

by G.A. Cooke

There is really very little to report, correspondence having been relatively light and principally consisting of notices from head office.

In December 1969, I received a letter asking if I, as Secretary for the chapter, wished to have the membership cards mailed to me or would I prefer that they be sent to the membership secretary. Since we do not have a membership secretary, I of course replied that the cards should continue coming to me. However, this raises a possibility - as the chapter grows, perhaps we should have a separate person as membership secretary who could welcome new members by a letter, something which I am sorry, I have not had time to do.

Another letter received was regarding several members who had not renewed their membership - would I write and if possible find out the reason for non-renewal. For most, it appeared that time flies by too quickly and, once reminded, renewed their membership. A few did not; the reasons given were mostly two-fold -1) 'I find my interests are not really those of the Society.' and 2) 'I was torn between rejoining Special Libraries Association or A.S.I.S. I decided for S.L.A. as being closer to my interests.' Thus the proposed merger between SLA and ASIS may solve these member's problems.

I could not possibly give a report without saying a special thank-you to Doreen Heaps and Janice Heyworth. Without their help this conference would not have got so well off the ground and I would not have survived the year as secretary-treasurer.

Turning to my duties as Treasurer - I am pleased to say we are in a healthy state financially, having some $450 in hand. This does not include all income and expenses from this conference, the bank balance having been ascertained just before leaving for Vancouver. A copy of the financial statement prepared is attached, together with the previous one made up to Dec. 31, 1969. This last was made to report to the parent organization, since they require statements of accounts from Jan. 1 to Dec. 31 in any one year.

In closing, I would like to ask for contributions to the newsletter. At present, I am volunteer editor for a newsletter issued quarterly to inform members of our chapter and also members of the Alberta Information Retrieval Association. I am fairly well supplied with news of activities in Alberta since I know many people active in information activities in the province. But there must be many provinces and territories of our chapter, and I urgently request you to send in contributions so that the newsletter truly represents Western Canada.

Not having read the constitution too thoroughly when accepting nomination last year as secretary-treasurer, I had not realized I was letting myself in for a two-year term. I will do my best in the upcoming year.

Financial Statement from the Western Canada Chapter of ASIS
Year 1969

Balance Jan. 1, 1969		$ 0.00
Income		
Chapter remittance Apr: U.S. $25 =	$26.78	
Nov: U.S. $39 =	$41.78	
Registration and membership fees collected at inaugural meeting (less exchange on checks)	$622.33	
Interest from bank: Oct:	5.02	
Total:	$695.91	$695.91
Total:		$695.91
Expenditures		
Re inaugural meeting:		
Printing	$ 76.10	
Facilities	$170.20	
Membership fees to ASIS	$231.64	
Bank charges	.15	
Total:	$478.09	$478.09
		$217.82
Total:		$695.91

Submitted January 19, 1970

(Prof H.S. Heaps, Chairman)

(Nita Cooke, Mrs., G. A.
Secretary-Treasurer)

FINANACIAL STATEMENT OF WESTERN CANADA CHAPTER OF ASIS

Balance January 1st, 1970		$ 217.82
Income:		
Bank interest April/70	$ 3.24	
Chapter remittance from ASIS April/70	57.01	
Registrations for annual meeting . . .	95.00	
	155.25	155.25
		$373.07
Expenses:		
Gestetner - paper for newsletter . . .	13.51	13.51
	13.51	
Bank balance Sept. 12/70	$359.56	359.56
		373.07

Outstanding expenses:

Printing of brochures	$ 71.25
Secretarial	28.00
Mailing of proceedings	16.00
	115.25

Therefore true balance is $257.82

THE MARC TAPES

P. Simmons
Chairman

Search Programs for MARC Tapes
at the
University of Alberta

by D. Heaps, V. Shapiro, D. Walker, and F. Appleyard*

This paper is a report on a MARC project carried out in the Department of Computing Science under the direction of Professor Doreen Heaps with the cooperation of the Library Systems Group, University of Alberta Library, and of Professor G. Pannu of the School of Library Science. In this project several students wrote experimental programs to manipulate the MARC tapes. The MARC (Machine Readable Catalog) tapes are subscribed to by the Library Systems Group. One month's supply was used in the experiments.

The experimentation was concerned with three aspects of computer manipulation. i) Programming to achieve fast code conversion from ASCII to EBCDIC and to perform work counts. ii) Programming to dump, strip and relate fields from the MARC tape. iii) Programming to reformat the MARC tapes to search them on author and title using the programs developed at the University of Alberta for searching Chemical Titles. A lecture summarizing the experiment was given by Messers Shapiro, Walker, and Appleyard to a joint meeting of students from the Department of Computing Science and the School of Library Science.

Code Conversion and Word Counts

The first program for the MARC tapes was written in Fortran with an assembler subroutine to do translation from ASCII to EBCDIC. The elapsed time for the program to run was over three minutes. Its sole purpose was to do a tape to tape translation.

In view of the poor time performance all programs were later written in assembler. In terms of time and space, assembler seemed to be more than five times as efficient as Fortran.

---

*Special thanks are extended to E. Bird of the Systems Development Group

As the MARC tapes received by Cameron Library are written in extended USASCII code, the simple translate program was needed to convert input records from the tape into EBCDIC. The records are variable length and unblocked with a maximum length of 2048 bytes. On this assumption the program TRANLT was written around the machine operation TR(translate), to translate from a given address the next 2048 bytes. As it does not change any of the registers the storing of the registers in the save area was omitted.

Many of the ASCII characters have no equivalent in the EBCDIC set. Thus untranslatable characters were arbitrarily translated into a vertical bar (upper case Y on the Model 29 keypunch). Because of the special meaning assigned to some of these characters information was lost, but for the purpose of the experiment the translation was quite sufficient.

As the labels on the tapes are also in ASCII they are indecipherable to the operating system. Insertion of the BLP(by-pass lable processing) parameter in the label field of the JCL will circumvent the problem of the operating system attempting to read and verify the labels. Any tapes read in this fashion should be notated as 'unlabeled 9 track' on the tape slip submitted with the job, as the operator must give special consideration to these tapes. The program ASCITOEB handles unlabeled, non-standard tapes. It was designed to read as many tapes as wanted--by concatenation of the JCL--putting the header and trailer labels out on the line printer and the intermediate data out in EBCDIC on another tape. The program was set up so that only the JCL need be changed to effect changes in the input and output devices. For instance ASCITOEB could be run as the first step in a job in which it reads in three tapes and leaves them as output on disk for the following steps.

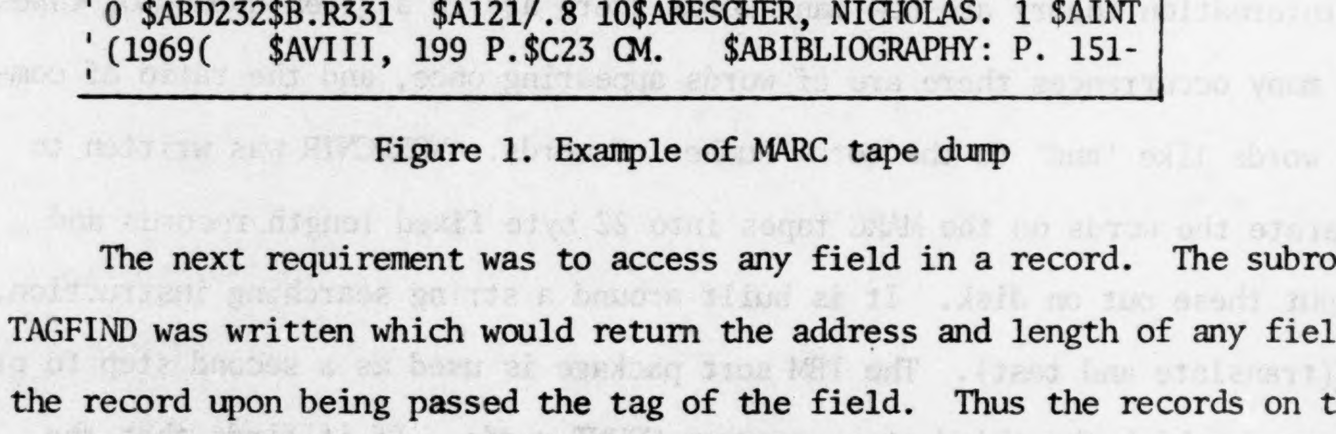
A problem arises when two or more MARC tapes are concatenated into one file by ASCITOEB. The original tapes had records ordered according to L.C. Card Number--a situation which no longer exists when these records are concatenated. The tape cannot be directly ordered by the IBM SORT/MERGE package as the L.C.Card Number is not at a fixed location within each record. Thus ORDER reads the tape, finds the position of the L.C. Card Number, duplicates the L.C. Card Number in a fixed eleven byte field immediately in front of the SCW(segement control word) and puts the new record out on disk as fixed block. The IBM SORT/MERGE procedure is now called as a second step, followed by a third step REAM. REAM reads the ordered output from the SORT phase and rebuilds variable blocked records.

A program which counts the number of words on the Chemical Titles (CT) tapes had been written by L. Thiel of the Department of Computing Science. A similar program was wanted and written for the MARC tapes. Items of interest to information theory are how many words there are in a given number of tapes, how many occurrences there are of words appearing once, and the ratio of common words like 'and' to the total number of words. WORDCNTR was written to separate the words on the MARC tapes into 22 byte fixed length records and to put these out on disk. It is built around a string searching instruction, TRT(translate and test). The IBM sort package is used as a second step to order the words which the third step, program COUNT reads. If it finds that the word that it is reading is the same as the last word it increments a counter, otherwise it outputs the word and count. From this base we can easily gather other relevant statistics.

The aim of this part of the project was to overcome the difficulties with I/O in assembler, associated with converting the raw MARC tapes into a form useable at the University of Alberta 360/67 installation. A following step could be taken in connection with this project. This would be the conversion of the record directory from EBCDIC to binary, a necessary step before large scale searching of the MARC data base is economically feasible.

Programming to Dump, Strip, and Relate Fields

A subroutine which is capable of handling all I/O for MARC oriented programs was written in assembler. Then a program was written which read the records off tape and wrote them on the line printer. This gave a dump of the tape in order to find out exactly what the information on the tapes looked like.



```
C0400NAM  2200145          00100130000000800410001305000160005408200110000
0214650001100244|   68009384 |690408S1969    NJU      B    00100 ENG
RODUCTION TO VALUE THEORY.|0 $AENGLEWOOD CLIFFS, N.J.,$BPRENTICE-HALL$(
186.|00$AWORTH.|
```

```
010000230008124500340010426000510013830000250018950400300
0 $ABD232$B.R331  $A121/.8 10$ARESCHER, NICHOLAS. 1 $AINT
(1969(   $AVIII, 199 P.$C23 CM.   $ABIBLIOGRAPHY: P. 151-
```

Figure 1. Example of MARC tape dump

The next requirement was to access any field in a record. The subroutine TAGFIND was written which would return the address and length of any field in the record upon being passed the tag of the field. Thus the records on the tape were capable of being "stripped" of fields of interest.

The next sequence of programs was concerned with the specific fields: L.C. CALL NUMBER and Topical SUBJECT HEADINGS. A program was written which read in the record, then stripped off these two fields and wrote them out. Thus, a sequential listing of the L.C. Call Numbers and Subject Headings was obtained.

```
P27.S68528  PR
FANTASY.

QK166  .C6
WILD FLOWERS MASSACHUSETTS MARTHA'S VINEYARD.

TE7  .H5 NO. 245
TRAVEL TIME (TRAFFIC ENGINEERING)

TL726.2.N6  N6 1967
AIRPORTS  NORTH CAROLINA  DIRECTORIES.

SB959  .P73
SILVEX.
HERBICIDES  TOXICOLOGY.
PESTICIDES AND WILDLIFE.
FISHES.
INVERTEBRATES.
FARM PONDS.

BM42  .S47
JUDAISM HISTORY ADDRESSES, ESSAYS, LECTURES.
JEWS HISTORY ADDRESSES, ESSAYS, LECTURES.
JEWISH LEARNING AND SCHOLARSHIP.
```

Figure 2. Sequential Stripping of L.C. Call Number and Subject Headings

Due to the interest in the relationship between these two fields a program was written which read the records, then stripped these two fields and wrote them onto disc. This file of L.C. Call Numbers and Subject Headings was then ordered alphabetically and written out. The ordering was done on both fields, that is the list could be written out with the subject headings ordered, or with the L.C. Call Numbers ordered.

```
CURRENCY QUESTION  CHINA.                         HG4572  C75 1968
CURRICULUM ENRICHMENT.                            LC3993  .R58
CURVES  JUVENILE LITERATURE.                      QA484  .R38
CURVES.                                           QA484  .R38
CYTOLOGY.                                         QH581  .B77
CZECHS IN THE UNITED STATES.    E184.B67          E184.B67  C29 1969
DAIRY PRODUCTS ADDRESSES, ESSAYS, LECTURES.       QP751  .S9
DAMS CALIFORNIA.                                  TC557.C2  A45
DEACONS.                                          BV680  .T5
DEACONS.                                          BX1912
DEAF  BALTIMORE.                                  HV2561.M3  F79
```

Figure 3. Stripping of L.C. Call Number and Subject Headings: alphabetical order of subject headings.

```
PIRATES.                                            G535   .G58 1968
AERONAUTICS  RUSSIA.                                G630.R8   B7 1968
AERONAUTICS  FLIGHTS.                               G630.R8   B7 1968
GEOGRAPHY  STATISTICAL METHODS.                     G74   .K47
STATISTICS.                                         HA29   .B53
STATISTICS.                                         HA29   .C5549
STATISTICS.                                         HA29   .C59
EDUCATIONAL STATISTICS.                             HA29   .E3 1969
MORALITY.                                           HB1481   .U55
ECONOMIC HISTORY  ADDRESSES, ESSAYS, LECTURES.      HB171   .S675 1969
ECONOMICS  ADDRESSES, ESSAYS, LECTURES.             HB171   .S675 1969
```

Figure 4. Stripping of L.C. Call Number and
Subject Headings: Ordered by L.C. Number

The writing out of such lists of specific fields can be very useful. A sample use, at a university, could be a list of title, author and subject headings of every book received at the library. This could then be ordered by subject headings and distributed to all local (departmental) libraries for their scrutiny. There is room in each record for added information; it could be used to indicate books available at the library.

The next problem was to link external L.C. Card Numbers with those present in the MARC records. A program was written which would read in any number of L.C. Card Numbers, them write these on disc and order them. The program then reads these ordered card numbers off disc, and the MARC records off tape and finds the matching record for each card number, if the record is on the tape. The call number and card number of each matched record are then written out. Thus, by inputting the card number the call number may be automatically retrieved. It is possible to write out any field of a matched record, not only the call number. It is, in fact, possible to write out the total information normally found on a 3 x 5 card, on card stock. This is the type of program that has been widely written elsewhere for MARC tape manipulation. In this project every effort was made to optimize the program within the limits of time available.

The general ability to strip and relate records of specific fields for varied purposes was emphasized in this part of the project.

SDI Search Programs

The third part of the project required the design of a retrieval system for the MARC tape data base. To do this one must have an information retrieval program which is suitable to the structure of the data base. It was decided to change the data base format rather than write a program to fit the MARC format because of the presence, at the University of Alberta, of highly developed retrieval programs for Chemical Titles.[1,2]

The Chemical Titles (C.T.) search program was chosen because it is a proven program and the conversion to Chemical Titles format does not present a serious difficulty.

The three main records of the C.T. format (author, title, page) had to be augmented to allow the full use of all information in each MARC entry. The final records of the converted MARC tape are classified as author, title, subject heading, abstract, reference and imprint. Each of these records is of either Author Type of Title Type (of which a brief description follows).

Author Record

An author record is of fixed length 81 characters.

The first 17 character segment is the reference to the article, if any.

The eighteenth and nineteenth characters are the record type and number in sequence respectively. The records are sequenced, one after the other, since one article might have more entries than can fit on one record and must then be extended. The type for author records is "1". The twenty-first through eightieth character are segmented into three, 20-character areas (21-40, 41-60, 61-80), each of which consist of the author's surname and initials. If an author's name is longer than 20 characters, the name is truncated from the right, to enable it to fit in its 20 character segment. Should there be more than three authors, then the next record would be the second author record in sequence and, therefore, would have a number sequence one greater than the previous one.

Title Record

The title record, which is of type "2", differs from the author record in that characters 21 through 80 contain the full title text. Should the title extend to more than 60 characters then the remaining title terms would overflow onto the next title record in sequence. No terms are truncated, so, should any term fall on the "boundary" between records the term is carried over to the next record. As was previously stated the reference of the title record is identical to that of the author record and consists of journal coden, volume, pages, etc.

Format Conversion

We have a program which will search a specified data base, but we have no data for it to search. Thus we come to the program designated to convert the MARC tape entries into records which are acceptable to the C.T. search program.

Some of the MARC records are described below in a manner suited to the conversion. For example, "extended titles" and "notes" are "abstracts".

The entries thought to be most improtant were: author, title, subject headings, abstract, L.C. Call Number, Dewey Decimal classification, collation and imprint.

The L.C., Dewey, and imprint were converted into one author entry as was the author entry in the MARC record, while all others are considered to be title records. A search question must be formulated to suit.

When converting the MARC data to C.T. format one does not use *all* information on the MARC tapes. In each variable field, which is the information content area, there are many different "field descriptors" which describe the various fields in the area. In most cases these descriptors were deleted since all content was deemed necessary with the exception of the author entry. As was shown before, the author record takes the surname and initials of the author. For this conversion it was thought adequate to enter only the author's surname.

There is one other major change from the regular C.T. format; there is no reference in columns 1 through 17. This deficiency comes about because most entries on the MARC tapes are of books and not of journals, that are referenced by CODEN, volume, page, etc. Now, for the C.T. program to function all records must be unique from all others on the tape. This can be very easily arranged by putting in a record counter and placing this number on each record.

Therefore, for each record, even though the number will still be the same, for each title, author record, etc., the type and number entries will allow for dissemination between all entries.

The C.T. program also requires a "last record" to function properly. A last record is one in which all entries finish on the same record. For this we chose the L.C. Numbers, Dewey Classification, and Imprint Record, which is "9A". It is very unlikely that a book (or other entry) would be totally void of all three of these entries. The "A" in the sequence number position mentioned above occurs, since the collating sequence of S/360 places letters before digits. It is quite possible that a record will need more than 10 records to complete the conversion, but if we start with "A" through "Z" and then continue with digits we can accommodate over 30 sequenced records of one type.

Figure 5 illustrates the reformatting of the MARC records.

```
MARC70     1A WILLIAMS
MARC70     2A WHAT'S IT ALL ABOUT? A NATURAL PHILOSOPHY FOR OUR TIMES,
MARC70     2B  GOERGE L. WILLIAMS
MARC70     4A PHILOSOPHY OF NATURE.
MARC70     4B COSMOLOGY.
MARC70     4C MAN.
MARC70     8A NEW YORK, EXPOSITION PRESS (1969)
MARC70     9A BD581 .W47             113/.2             177 P. ILLUS. 22
MARC70     1A GOLDFARB
MARC70     2A A TIME TO HEAL; CORRECTIVE SOCIALIZATION: A TREATMENT
MARC70     2B  APPROACH TO CHILDHOOD SCHIZOPHRENIA, BY WILLIAM GOLDFARB
MARC70     2C  IRVING MINTZ, AND KATHERINE W. STROOCK.
MARC70     4A SCHIZOPHRENIA.
MARC70     4B CHILD PSYCHOTHERAPY RESIDENTIAL TREATMENT.
MARC70     8A NEW YORK, INTERNATIONAL UNIVERSITIES PRESS (C1969(
MARC70     9A RJ506.S3 G6            681.92/8/982        1X, 148 P. 23 CM.
```

Figure 5. MARC records reformatted for Chemical Titles search

Figure 6 illustrates searches resulting from use of the C.T. search programs on the reformatted data.

```
I          MARC TAPE CONVERSION

      AGAIN    HEAPS
      AND TL CANAD
      AND TL LIBRAR

      CAMPBELL
      CANADIAN LIBRARIES (BY) H. C. CAMPBELL.
      LIBRARIES CANADA.
      (HAMDEN, CONN.) ARCHON BOOKS (1969)
      Z735 .A1C29 1969B   021/.00971         90 P. 23 CM.

      THERE WERE          1 TITLES FOUND FOR THIS QUESTION

II         MARC TAPES CONVERSION

      AGAIN    HEAPS
      AND TL LITERA
      AND TL CRITIC

      FRYE
      LITERARY REVIEWS AND CRITICISMS.
      LITERATURE, MODERN ADDRESSES, ESSAYS, LECTURES.
      NEW YORK, GORDIAN PRESS, 1968.
      PN710 .F77 1968B    809                VIII, 312 P. 23 CM.

      THERE WERE          1 TITLES FOUND FOR THIS QUESTION.
```

Figure 6. Two Searches on MARC tapes

The search programs were written in PL/I.

One must remember when using the conversion program that, while it is set up to convert the aforementioned entries in each record to C.T. format, it will convert any field in a MARC record which may be required, and thus one will be able to search on any required data. Documentation for these programs is available from the Department of Computing Science, University of Alberta. Some sections of this work were supported through NRC Grant A5250. The programs were written during the period December 1969 to April 1970.

REFERENCES

1. Thiel, Larry H., and Heaps, H.S., Computer Search of Chemical Titles at the University of Alberta, Department of Computing Science Publication No. 15, Oct., 1968.

2. Heaps, H.S. and Thiel, L.H., Optimum procedures for economic information retrieval, Information Storage and Retrieval, 6, pp. 137-153, 1970.

3. Information Systems Office, Library of Congress, Marc Manuals Used by the Library of Congress, Information Science and Automation Division, American Library Association, Chicago, 1969.

4. Information Systems Office, Library of Congress, Subscribers Guide to the Marc Distribution Service, Washington, 1968.

5. Information Systems Office Project Marc, An experiment in automating Library of Congress catalog data, Library of Congress, Washington, 1967.

6. Auram, H.D., Knapp, J.F., and Rather, L.J., The Marc II Format; a communications format for bibliographic data, Information Systems Office, Library of Congress, Washington, 1967.

7. Council on Library Resources Inc., Proceedings of the Fourth Conference on Machine-Readable Catalog Copy, Library of Congress, Dec. 4, 1967, Library of Congress, Washington, 1968.

8. Information Systems Office, Library of Congress, The Marc Pilot Experience; An information summary, Library of Congress, Washington, 1968.

9. Washington State Library, Marc, ALA Bulletin, Book Catalog of King County Library System, North Central Regional Library, Timberland Library Demonstration, 1967.

MARC
at The University of Saskatchewan
Saskatoon

A.S.I.S. Presentation (1970)

Western Canada Chapter

by Grover C. Burgis

and

Edwin Buchinski

MARC AT THE UNIVERSITY OF SASKATCHEWAN

This is not an academic treatise, nor a learned paper, neither is it a new theory on information sciences. Rather, I have been asked to present a paper quite simply 'How we done it! at Saskatoon.'

Why MARC?

When I arrived at the University of Saskatchewan Library, I was faced with two major problems, (a) faculty dissatisfaction with the slow processing of newly acquired items, and (b) rising costs of technical processing. Kilgour from Ohio says "The per-student costs of libraries are rising somewhat more than twice as rapidly as unit-costs in the general economy. Only the introduction of an increasingly productive library technology can reduce the rate of rising costs. Here the only apparent, fruitful avenue of technology is that of the computer employed as an information processing machine."[1]

September 1967 was an opportune time for the University of Saskatchewan Library to start automation on a central bibliographic file. No such mechanized system had been started to date, whereas other University libraries had started experimenting with MARC I or similar formats, in batch processing modes on second generation computers, like the 7040. We at the University of Saskatchewan were able to start outright on an IBM 360/50. My point of view was to start work on a total system concept for the library and therefore, my aim was to capture bibliographic data in acquisitions to start building what Warheit[2] calls a central bibliographic file, which can handle a variety of library applications more economically than a separate set of files for each job, i.e. Acquisitions, Cataloguing, Circulation, production of a union catalogue, etc.

Another influence as to Why MARC? was my experience at the University of Pittsburgh. I had just arrived from the United States and was pro MARC as a North American Standard, and felt that a MARC type file would naturally form the basis for such a central bibliographic file.

In addition, library professionals were doing L.C. editing, and although we already had camera equipment no regular photographic routines had been established for making copies of NUC entries. At that time, I decided that if the manual system had not been perfected over the last 10 years to a satisfactory state, then something more than a time and motion study was needed, and therefore, decided to get MARC printouts into the hands of the Cataloguing Department as soon as possible.

I'm very much aware that many people at this seminar would find just as many reasons for not going to MARC, but I hope these preliminary comments will help to explain my reasons for doing so.

Computer Environment

It is inevitable that every library's automation program will be constrained by the computer resources at its disposal. Libraries with large computer resources can undertake more sophisticated applications. We at the U. of S. do not have our own computer, but use the facilities of the Central Computer Centre. Figures 1 and 2 have been selected to give you some appreciation of the constraints that our systems people at the Computer Centre work under. Our facility comprises an average IBM 360/50 installation. We have 3 nine track tape drives at our disposal. The core consists of 384K and this is being considered for expansion in a year or two. At that time, we will probably have to make major additions to the peripheral equipment, and acquire a larger computer model.

We work under the MFT° II / HASP+ II versions of 360 operating system. Previously, we worked under a total execution system, where the CPU was occupied from input to output.

Occasionally, our core presents a processing constraint because of the way it is partitioned. 60K is devoted to the HASP program which resides in core. 20K is assigned to the circulation program which runs our on-line CRT terminal for the library's circulation records. There are two user partitions consisting of 100 and 150K. The latter two can be combined for large programs when necessary. Our remaining 54K house the operating systems software. All the library programming has been done in COBOL, even though Fortran, Assembler, Algol, Snobol, and PL 1 are also available.

MARC Programmes - Used at the Library, University of Saskatchewan, Saskatoon.

The following programmes have been written and are used to run the MARC tapes at our library.

(i) a program to translate the tapes from ASCII (American Standard Code for Information Interchange) to upper and lower case EBCDIC.

* (ii) a program to translate the tapes from ASCII to upper case EBCDIC.

*(iii) a program which uses the weekly MARC tapes to update the Current MARC file.

° Multiprogramming with a fixed number of tasks.

+ Houston Automatic Spooling and Priority System.

* No longer used.

* (iv) a program which produces a formatted print of selected MARC records.

(v) a program to print catalog cards from selected MARC records.

(vi) a program to list MARC records in call number sequence within subject groupings.

(vii) a program to dump a MARC record.

(viii) a program to transfer selected records from the Current MARC file to the In Process file.

(ix) a program to update the Current MARC tape and to produce author-title codes from the new weekly MARC tape.

(x) a program to update the new code tape with codes produced from each new weekly MARC tape.

(xi) a program to split off old records from the Current MARC tape and to produce author-title codes for these old records.

(xii) a program to remove the old codes from the new code tape.

(xiii) a program to produce author-title compression codes from unverified author-title requests.

(xiv) a program to list the author-title requests and codes with the corresponding L.C. card numbers of MARC records matching the codes from the requests.

(xv) a program to update the history tape with records split from the Current MARC tape.

[Standard Book Numbers are also produced as access points in the above programmes. (ix - xiv)]

* No longer used

Figure 3 illustrates a MARC mini tape as compared with its standard 2400 foot computer counterpart. We receive the 9 track 800 bpi weekly MARC tapes, and these have contained from a low of 310 records to a high of 2086 records for a single week. One large computer tape can contain approximately 66,658 MARC entries. Therefore, a book collection of about 300,000 titles would require five tapes for record storage. A weekly numeric listing accompanies each MARC tape and simply lists the L.C. numbers on that tape for all the records which it contains. A sample page of this listing is illustrated in figure 4. Figure 5 shows a dump of the MARC tapes in which the various data elements have been identified.

In this brief section of our presentation I will attempt to make you aware of the MARC tape format and to also illustrate the need for its sophistication.

As most of you are no doubt aware, the MARC format while spearheaded by L.C. also embodies the recommendations from librarians who worked with machine-readable data before MARC was a reality. Before and since its inception, MARC has been subject to changes as new applications and needs are realized. This metamorphosis is possible and can be expected to continue as MARC is modular in design. Some of the illustrations which have been used suffer as a consequence of the changes made to MARC, but I hope you will bear with us in this cursory examination and refer to the fourth edition of the MARC manual for most of the specifics which have been overlooked because of the time limitation.

Prior to showing you a sample MARC II record, I thought that you might appreciate knowing a little about the method by which this machine-readable catalogue is produced. Illustration 6 is of a sample worksheet that is used by LC to produce LC cards. This worksheet was used prior to the initiation of the MARC project, and is still being used to produce LC cards now that MARC is operational.

Because the computer requires that the information which existed implicitly on the LC card be made explicit, LC had to devise a system of coding and a worksheet on which to represent these codes. You can look at the main entry on an LC card and know that it is the main entry, but the computer can't recognize this information as such. Unlike a human, the computer requires that each data element be identified or made explicit.

The traditional worksheet, after being completely edited for descriptive and subject cataloguing errors, is used to produce LC cards. Prior to being used for card production purposes, this worksheet is matched with a clear plastic overlay that contains spaces for codes needed to make the LC card information explicit to the computer. The original worksheet and the overlay are then xeroxed together to produce the MARC Input Worksheet.* Notice that the new worksheet differs from the previous slide only in the following respects. First, it has room for inserting codes (tags) which identify parts (fields) on the traditional LC card, and second, it has a matrix for fixed field data elements.

The next step in the MARC record production routine consists of passing the unedited MARC input worksheet to an editor who is familiar with the various fields of information on an LC card, and who knows the LC tagging or coding scheme. This person is responsible for assigning mnemonic tags which will identify the fields of the bibliographic data for the computer.

No editing is required for the descriptive and subject cataloguing that appears on the MARC Input Worksheet. The editor, however, will cross out those elements of information which have already been identified for computer purposes by the mnemonic tags, for example '1. Title' and 'Series' in the tracing information represented on the sample edited worksheet.

---

* See figure 7

The editor will fill in the fixed field information and also insert the delimiters. Notice in illustration 8 the crosses or delimiters that appear within the imprint field. These separate the data elements within that field. Following the editing, or the assignment of tags and codes, the MARC Input Worksheet is passed on to the person responsible for transferring the worksheet information to a medium that can be read by the computer. Currently, an IBM MT/ST is being used by the Library of Congress to enter the worksheet information on a typewriter which in turn produces a hardcopy[*] and a magnetic tape. The hardcopy shows the typist exactly what she has transferred to the tape which the magnetic tape serves as the computer input medium.

After feeding the magnetic tape input into a computer, the Library of Congress receives a diagnostic printout.[+] This printout represents all the bibliographic information that will be recorded on the magnetic tape which will be distributed to those libraries subscribing to MARC. The diagnostic printout exhibits all of the information that was shown in the previous illustration, plus machine inserted information such as indicators and delimiter codes. The latter have been supplied as a result of computer interpretation of the implicit and explicit coding done on the MARC Input Worksheet. (see figure 8)

The above illustrations were produced from handouts which I obtained at the Seattle MARC Seminar in 1968. To my knowledge, these routines are still being used today, but perhaps Mr. Simmons might have more recent information which would contradict what I have said.

---

*See figure 9

+See figure 10

The Library of Congress has defined format as consisting of three basic elements. According to Paul Reimer, these elements are "structure, designators and content. Structure is the physical representation on tape, capable of containing the bibliographic description for all forms of material; content designators are the labels to identify explicit data elements for particular material; the content is the data itself."

This will be the standard method in which machine readable information will be structured on MARC tape regardless of whether it represents serials, maps, motion pictures and filmstrips or some other form of publication. In the following illustration* we can see the first of Reimer's three elements, the structure as being represented by the leader, the record directory and the variable field data.

The leader contains information pertinent to the entire record. Its first five digits are used to represent the length of the entire record in terms of the number of characters which are contained within the record, starting with the first digit of the leader and extending to the record terminator symbol. A record status symbol is used to inform the computer on whether that particular record is new or if it represents a correction or whether it should be deleted from your machine-readable file. The "a" informs the computer that this machine-readable record contains bibliographic information about a unit of printed language material and the "m" denotes that this bibliographic unit is a monograph.

---

*See figure 5

Following the two blanks two single characters inform the computer of the number of positions in front of each field and in front of each subfield that are occupied by the indicator and the delimiter code respectively. The "base address of data" provides information on how many positions are taken up within the record by the leader and the record directory data, before the variable field information starts. The eighteenth position represents one of the most recent innovations to the MARC record. It will be used to indicate the degree of completeness of the machine record. A blank and a "1" code will indicate whether or not the machine record was produced from a physical inspection of the item the record represents. This is one outcome of the RECON study.

Now, let us look at Reimer's second basic element of the MARC format, namely the content designators. This element consists primarily of the record directory and to a more finite extent, the field indicators and the delimiter codes of the subfields.

The record directory contains information about the variable field data. Each field of information on an L.C. card, i.e. the main entry, title, collation, etc. is identified by one record directory entry. A record directory entry is made up of the tag, (3 characters) the field length (4 characters), and the starting character portion (5 characters). This tag identifies the field to which the entry refers, the field length gives the length of the variable field that is identified, and the starting character position informs the computer where that variable field data begins relative to the starting position of the variable field data, in the entire record.

For examples of indicators we have to go to the front of the variable fields themselves. The first indicator in front of the main entry informs the computer that the personal name begins with a surname while the second indicator informs the computer that the main entry is not the subject of the work. The delimiter codes[+] identify the subfields within a given field. #a on the collation identifies the place of publication, #b isolates the publisher and #c specifies the date of publication.

We might generalize and say that the third element or content within the MARC format is everything within the variable field data except the indicators, subfield codes and possibly the fixed field data. Fixed field information is used to make such additional items as the date of publication, the presence of an index, the language, etc. explicit about the bibliographic units that are being described in the MARC record.

The last two illustrations (11 + 12) in this section have been made up to give you some appreciation for the extensive number of tags and indicators that are used by MARC, and also to provide you with an impression of the many subfield codes deemed necessary to properly identify a machine-readable record.

In summation, I would like to stress that printing, catalogue division, information retrieval, and filing were some of the criteria that were used to judge the flexibility and usefulness of the MARC format.[*] The next illustration will show the flexibility or sophistication that can be achieved in filing. Regular computer filing would sort these three names according to the data elements a, b, c. However, library filing rules stipulate that these names be sorted in the acb name element sequence. If subfield codes hadn't been assigned then the computer would be able to file these names only in reverse order to that which librarians prefer.

[*] See figure 13

[+] Note that the delimiters, the field terminators and the record terminator are all non-printing characters which do not appear in the dump. (illustration 5)

Figure 14 shows how three different tags can be used to divide one name among two different catalogs on the basis of the identification provided by the tag. The 100 and 700 tags for Smith would mean that, that information would be used in an author-title catalog while 600 tags are used to indicate subject headings which would print in a subject catalog. No illustration was provided for a sample of the usefulness of the MARC format for information retrieval. However, it is easy to see that because the main entry has been identified, the task of retrieving information on all the books written solely by Smith, John, can be restricted to checking the main entry of all records for books with a main entry tag of 100, in which the first indicator is 1, and only then on the basis of a letter by letter match for Smith, John. The information retrieval application would mean that a letter for letter match on the main entry would be performed only on records for books with a personal name, surname first, instead of the entire data. However, the ultimate benefits to be derived from the tagging scheme for information retrieval purposes has yet to be exploited.

The last sample of MARC flexibility for printing purposes will be illustrated by the differences in formatting of 3 x 5 cards later on in this presentation.

Unit Card Printouts

As I mentioned in the beginning, my original aim was to get MARC copy into the cataloguing department as quickly as possible. Therefore, I'll show you how we have printed out the records from the beginning through the various stages of development.

Note that the first four styles of the printouts are in upper case only. It was not until some time in February, 1970 that we rented a TN print chain which gives us upper and lower case but still is limited in the number of characters that it has. We still cannot process French or Italian languages because the diacritical marks are unavailable on this chain. You will notice in the illustrations, that a blank is used to alert the MARC editor, whenever a diacritical mark is necessary. This illustration (15) shows the first printout we made from the MARC tapes in May 1, 1969. We never actually used this printout in the cataloguing department, but immediately revised it to an improved format. (see figure 16) The cataloguers did edit the latter printout for several months. Our third style of printout was formatted for a 3 x 5 card and printed on standard computer paper stock.* The fourth printout produced the MARC catalogue data on actual card stock.* Our current style (see figure 19) is in upper and lower case. These printouts are edited with the book in hand, and then go directly to the typist in the typing pool to produce multilith masters for the card set production (see figure 20) You will notice black checkmarks on the illustrations which must be eliminated before we start producing card sets. Several reasons have been given for the occurrence of these black checkmarks. One reason being the thickness of the card stock. It has also been suggested that the speed of the print chain may be too slow, and that the adjacent character drags. Other suggestions have been printing density, hammer speed, print chain construction, gold slug position, and hammer striking position.

I should like to remind you at this time that LC and ALA have approved a 174 character set print chain in upper and lower case with full diacritical marks. I saw the overlays for this print chain when I was in Washington in March, and the print font looks similar to the style on the LC printed cards. As soon as this is

*See figure 17

*See figure 18

available as a shelf item from a computer manufacturer, we will rent it and then be able to print foreign languages with roman alphabets.

Catalogue card format

The MARC format is extremely flexible and allows a user library to design a catalogue card which is compatible with its present style of cards, or one that might be more practical for computer formatting. For example, the Library of Congress formats its class number to print on the first line, with any decimal portion following on the second line. To conform to the University of Saskatchewan's formatting of the call number, we have put the alphabetic and numeric portions of the class number on separate lines. Some of the additional conventions which we use are: the omission of a period before the cutter line; spacing of the main entry and the title paragraph as outlined by Bidlack[3]; a two spaced indention is used whenever a subject, a series, a main or an added entry heading continues for more than one line; no blank line is left between the title paragraph and the collation statement; two spaces to separate the elements in the collation; a recent revision will provide two spaces between the fields in the title paragraph; and bibliographic notes are printed in the sequence that they appear on the MARC tapes.

This illustration (21) shows a catalogue card from NELINET,* a MARC user in Boston. I have included this to emphasize the flexibility of MARC card printouts. The latter supplies a period before the cutter line, uses two dashes to separate the subject heading from its subdivision, provides only a single space between the elements in the collation statement, and abbreviates "Title" in the tracing to "T".

* New England Library Network

Figure 22 provides an example of the statistics that are produced at the end of every MARC run. These are essential for evaluating our MARC usage. The present technique lists the number of requests submitted, the number of MARC records found, and the number not found. All requested L.C. numbers which are not on the MARC tape are listed in ascending L.C. card number order, and are retained for future submission against the updated MARC history tape.

The next illustration (23) is a sample of our cataloguer's edit sheet. This edit sheet consists of a unit card and a diagnostic. This form is now finalized and is ready for use when the Acquisitions/Cataloguing system goes into effect later this fall. The purpose of this edit sheet is to enable professional cataloguers, with the book in hand, to revise any data element on the catalogue card portion which they feel is incorrect. For instance, if the main entry should be spelled Ottoburg Show,+ the cataloguer would simply correct the main entry field in the diagnostic. The revision would then go to the keyboarders who would rekey that main entry for this particular record, and the following day the cataloguer would receive a revised edit sheet. If the latest revision meets with the cataloguer's approval, he may then submit the 'cataloguing edit complete' status card for this record, signalling that a complete set of cards should now be produced. Thus it means that the cataloguers will have the final say and final approval of all cataloguing. The cataloguer's edit sheet for books which require original cataloguing, will contain a printout of all bibliographic data that was captured at the point of acquisitions, although this will not be as complete a record as it will be in the case of books with MARC supplied source data.

+ See figure 23

Everyone will no doubt be interested in our MARC usage. However, it is extremely difficult to indicate in any meaningful way what this utilization has been. Therefore, I have drawn up some charts to give you an idea of how we have used the MARC tapes. Figure 24 indicates the total monthly MARC usage from July 1, 1969 through to the end of April, 1970. The top of the bar graph indicates the total number of unit cards requested, and the white portion indicates the total number of printouts received for any one month. You will notice that April, 1970, with 747, is the highest month for printouts so far. A fairly steady increase in the number of printouts is indicated, with an average of 390 per month, for the period July 1/69 to April 30/70. For the past four months this average has risen to 515 per month.

Total monthly MARC usage by application up to April 1970 is illustrated in figure 25 . The top of the graph indicates the total number of printouts received. That is, the total number of unit cards printed out for April 1970 was 747. You will notice here the use of the terms pre and post acquisitions. In general, pre acquisitions usage refers to MARC printouts that have first been used for ordering purposes. Post acquisitions statistics indicate MARC printouts which have definitely been used for cataloguing purposes only, whether the actual request was submitted by the acquisitions department or the cataloguing department. The acquisitions department requests MARC printouts for all books that lack LC copy at the time the book is received by the library. Cataloguers can also request these printouts for books when the acquisitions department has failed to locate LC copy. The latter items are usually from the cataloguing backlog of our work load. Therefore, the dotted area of this graph which represents the pre acquisition of an item, plus the white bar graph indicating MARC usage for cataloguing purposes, is equal to the total number of printouts for the month, as indicated by the top of the bar graph. No attempt has been made to manually count the number of pre acquisition printouts which have been utilized for post acquisition or cataloguing purposes. Such statistics would provide even a better idea of the MARC tape impact on cataloguing, but we don't feel that ascertaining this figure is essential at the present time, for it will be obtained much more easily once TESA I‡ is operational.

Figure 26 indicates our monthly post acquisition use of MARC. The top of this graph indicates the total number of titles received by our acquisitions department. The top of the bar with the cross lines, indicates the predicted number of English language monographs that were received. We estimate that 80% of our current buying is in English language monographs. The white portion of this bar graph gives the number of MARC printouts which were used for titles received. A total of 518, in April, 1970. However, if phase 3* of the RECON project had been a reality, we presumably could have received 1734 MARC printouts, or 80% of all the titles bought in April. Conversion of all English language monographs as the third priority* in the RECON project, should be realized by the time the RECON project is 50 per cent finished. We can expect to receive MARC printouts for 80% of all our book receipts at that time.

Figure 27 represents the monthly pre acquisition use of MARC. This graph follows the logic presented in figure 26 , however it was developed to provide you with an idea of the number of MARC printouts that are used to supply verified data at the book ordering stage. MARC usage in this area shows a steady increase, and will continue to do so as RECON progresses. On the other hand, post acquisition usage should eventually be confined to requests for very recent publications only.

---

‡ Technical Services Automation Phase I.

* Third category is 1897 - 1959 English language monographs
First category is 1960 - present English language monographs

Acquisitions

We are also working on an automated acquisitions system. Two basic problems have loomed in my mind from the beginning of our work. a) How to access a machine readable in process file in the acquisitions department alphabetically, At the present time we have a manual card file by author, which is not too reliable. A main entry approach is not reliable either at the acquisition point. At the University of Pittsburgh there was much discussion about the main entry concept, and the need for another access method for machine readable files. Therefore, I began to inquire about an author title approach to this problem.

b) The second problem soon arose when we started using the MARC tapes. How do you access the MARC tapes by other means than the LC card number, or SBN. Eventually we began to experiment with author/title compression codes and we are hoping that this method of access will overcome the above two problems.

Now I'll let Ed talk to you about the experimental work he and Bill Newman have done in this field.

From Mr. Burgis' presentation thus far, you are aware that access to MARC tape bibliographic data is restricted to the L.C. card number approach. Use of this data is impossible whenever the L.C. card number is not available unless the MARC subscriber has developed an alternate access point. I am not aware of any recent communications from the Library of Congress which would contradict the initial guideline that exploitation of the MARC data is the onus of the individual subscriber. Some hope was expressed at the March MARC users conference in Washington that author-title indexes might be made available by L.C. if ever that organization should choose to print book catalogs in the register and index style used by the Washington State Library.

At the University of Saskatchewan, the main reason for wanting to capture MARC information at an early stage in the automated system is to have accurate and complete bibliographic information for ordering purposes, which in turn could be passed along to cataloguing. Formerly, the L.C. card number limited our exploitation of this data base in that it was seldom available on an order request from faculty. Unless verifiers located an L.C. card number in the N.U.C., C.B.I. or some alternate source, MARC data which might have been on the tapes was unavailable for library use. This restriction had larger economic implications in an automated system since it means that in house keying of inaccessible machine readable bibliographic data would be required.

"A comparison of arrival dates of L.C. proofslips and corresponding MARC magnetic tape records at the University of Chicago revealed that four-fifths of the MARC records were received the same week as, or earlier than, the proofslips".[4] If full advantage of the prompt service provided by the MARC tapes is to be realized, it seems essential that author-title access to MARC data must be available as it is in the proofslip service. Mr. Payne at the University of Chicago was recently quoted as saying that Chicago is currently working on methods to broaden the scope of MARC matching by printing out MARC arrival cards for use in order selection; they will replace L.C. proofslips.

At the University of Saskatchewan, Saskatoon, we felt that our access point or author-title code should possess the following qualities:

1. It should minimize the effect of spelling errors in the data for which verified bibliographic information is sought. A misspelled request will fail to provide an accurate match. The amount of variance of the unverified data with its verified counterpart can be insignificant in a manual operation, but this small variation is significant in a computer application. An omitted mark of punctuation, inaccurate spacing or a truncated element in the submitted data, will produce a false hit or no match with its verified counterpart record, unless programming is undertaken to modify these discrepancies between verified and unverified information.

2. The request for verified information can be made in a few significant terms. In the proofslip service, these terms are the main entry and title. At the University of Saskatchewan, requests will be made using the surname of one of the authors of the book, and the first four significant words of the title. Initial articles and words with only one consonant will be disregarded, and our search request will not be restricted to the main entry.

3. The two previously stated requirements can be easily incorporated into a computer program to generate access codes from both verified and unverified data.

4. The resultant codes are short, of specified length and easy to search. An index made up of access codes is much more economical

to search than would be a data base consisting of a specified number of characters from the author and title fields.

5. Ideally the author title codes should be unique. Each code should refer to only one bibliographic record if it is to provide a key to bibliographic records stored on direct access devices.

An alternative index to the MARC tapes that might be considered is an author-title printout. If on September 9th, we had produced a cumulative author-title listing, then it would have consisted of 16,476 titles. All the MARC records since the first week in June would have to be printed, since the latest NUC catalog in our library is May, 1970 supplement.

Allowing one line per title, which might prove very inadequate, and 54 titles per page, such a MARC listing would be 306 pages long. This document would frustrate the user through its bulk, the document's size would limit access to this truncated bibliographic record, and the costs of printing and updating would probably be high.

In deciding to adopt the author-title compression code approach, the University of Saskatchewan, Saskatoon Library, has been persuaded by a number of advocates of such an access method. Probably the foremost investigation of the problems associated with bibliographic retrieval using input data which has varying accuracy is the one conducted by Ruecking of Rice University. Kilgour of Ohio University, Nugent of Inforonics, and University of Chicago and the Library of Congress are also enthusiastic about the possibilities of such an access method. It must be stressed that the theory of author-title codes is the same even though the mechanics of producing these codes differs among the various proponents of this system.

Our initial experiment using a small file revealed that our code in comparison to Ruecking's, produced fewer false drops, is easier to construct, is shorter and has the same retrieval performance. It also produces unique codes than would Kilgour's methods.

Figure 28 illustrates some of the various algorithms that have been used to compress the selected words from the author and title statements. Ruecking used a technique which produced four-four character abbreviations for both the author and the title statements. Kilgour used a code which consisted of three initial characters from the author's surname, the three initial characters from the title's first word and one character from the title's second word. At the University of Saskatchewan, we use the first 2 consonants of the first four significant words in the title and up to a maximum of four 2 consonant compressions for the first four words of the corporate author's name. Personal surnames are compressed to a 6 character code.

If you would please refer to the following flowchart, then my explanation of our planned and operational applications for the author-title codes might prove clearer. (see figure 29). In the "Weekly MARC Tape Processing" flowchart, you will notice that we produce author/title codes for each new MARC tape that arrives. These codes are merged with codes for the entire current MARC tape which consists of the current year's L.C. cataloging. The next photo (see figure 30) illustrates the use of MARC tape codes to obtain data from the MARC tape for pre-acquisition application. We submit requests for MARC data using only the information provided by the person requesting the order. The author and up to a maximum of four significant words from the short title are keypunched from the unverified orders onto IBM cards. These requests are submitted to the Computation Center for code generation and matching against the new code tape. The following morning we receive a unit card printout for each located entry and a listing of the entry/title requests which is used to match the unit card printouts with the original requests. The next two illustrations (see figures 31 & 32) provide a sample of the author/title, the SBN and the series statement access which we have developed at the University of Saskatchewan. No further details will be given on this application at this time since Bill Newman and I hope to provide these in a paper which we are submitting for publication. This application speeds up the verification process and in the automated system it will reduce the keying operations required in order placing and expedite speedy cataloguing for all the items found on the MARC tape.

At the MARC User Conference, Mr. Payne spoke of the problem of matching of the MARC data with a record from the in-process file and the difficulties of getting the MARC data where and when it is needed. Our approach to this difficulty is illustrated in the flowchart on "Post Acquisition Use of MARC", (see figure 33). We will produce author/title codes for all the records in our in-process file. Whenever new records are added to the in-process file, its codes will be matched, using a direct access technique, against existing codes to check for duplicate in-process records. This application should make duplicate error checking an exception rather than a routine task in the acquisitions department. The codes for the in-process file will also be checked against the codes for each new weekly MARC tape. Whenever a match results between the codes from these two files, then we may have obtained verified MARC data for an unverified record in our in-process file. A manual verification will eliminate any false hits and the resultant application should minimize the amount of manual searching for verified data after a book is received by the library.

The MARC Input into TESA I flowchart gives a systems illustration of how we will add MARC records to the in-process file.[+] In this series of flowcharts, which have been provided by Bill Newman, our system analyst programmer, the final one illustrates the update routine that will be used in the in-process file of TESA I.* (see figure 35)

Thus far I have avoided speaking about the problems of author-title code utilization. The primary difficulty in this application is that 100% retrieval cannot be obtained, mainly because of inaccurate unverified request data. However, neither can anyone contend that manual searches are 100% accurate. Secondly, false drops will result and therefore manual verification will be required to determine that the hits are actually desired. The following illustration (see figure 36) shows just how small a variation will produce a false drop. In the first example the two records differ in imprint only. The second illustration shows a difference only after the fourth line of the title statement. False drops however, can be a blessing in that you do retrieve very similar items which are parts of a series and some of these might be desired purchases. Government documents and corporate publications pose a high probability of producing duplicate codes. We hope to minimize this problem by submitting the personal author name whenever possible. The third drawback of the author-title codes, as planned at the U. of S. is that our technique does not possess the flexibility of an inverted file retrieval method, nor does it use the weighted search strategy employed by Ruecking.[5] In information retrieval terminology, the latter two techniques might verylikely increase our recall potential at the cost of reducing our relevance level. We feel that only experience with these codes can tell us what our most effective strategy will be in these projected applications.

In conclusion, we feel that the advantages of the author-title codes

+See figure 34
* Technical Services Automation Phase I

far exceed their limitations and the advent of on-line applications will definitely require such access methods. The active work that the aforementioned authorities in library automation have done, would concur with this view.

Costs

Everyone wants to know what our costs have been. Unfortunately, many people fail to realize that $100.00 spent at the University of Saskatchewan does not necessarily have the same purchase value elsewhere. The computer facilities we can buy for $100.00 are not necessarily equal to those that you can buy. If you can calculate the factor, or percentage that your costs are above those at the University of Saskatchewan, then perhaps you could estimate what your costs might be for similar development.

However, since there isn't any other way to indicate costs to you other than in dollars, we have charted a graph that might be useful as a guideline. (see figure 37)

Development costs consist of programming, keypunching and testing of the MARC tape programs. They also include maintenance costs for the developed programs, as the latter have required minor alteration to accommodate changes requested by the library. For example, the MARC printout formats have been revised three times, the method of requesting printouts and the statistics concerning MARC runs have also required changes. Minor alterations in the MARC tape records have been made by the Library of Congress and these have had to be accommodated into our programs.

In terms of development, we had a choice to make. Do we develop until we have a perfect card set, or can we utilize the printouts as we go? It was an arbitrary decision, and I decided to use the tapes as development proceeded. The development costs include three styles of printouts plus the cost of a high usage study.*

Update costs to May 1, 1970 included only the costs of translating the weekly MARC update tape from the ASCII code to its EBCDIC counterpart, and the costs of updating the current MARC tape (sort & merge). It includes $115/month rental for the TN print chain as of the middle of February 1970. This figure does not include tape rental charges. In mid February we stopped producing the upper case translation and history file. Therefore, update costs were cut considerably at that time. Since August the update costs also include the author/title code generation expenses as well as the updates to the new code tape.

Printing costs include the costs of producing MARC tape printouts for requested entries by the various library departments.

As of March 1970, our Computer Centre initiated a more sophisticated charging system. Formerly, it charged $60.00 per hour for total execution time. That is, spool to core through to output from core onto disk or tape. No charge was made for printing. Listed below are internal charging rates for 1969, indicating the new charging system.

Internal Charging Rates

Resource	Rate
CPU time	$52.00/hour
Core utilization	$ .10/K/hour

*see page 29

Internal Charging Rates (cont'd.)

Resource	Rate
Card reading	$ .36/1000 cards
I/O operations (tape & disk)	$ .52/1000 operations
Printing	$ .60/1000 lines
Card punching	$ 3.20/1000 cards

You will notice that our costs are down considerably for March and April.* The new charging system allowed us to take advantage of its various components and to reduce our costs. A significant cost factor under the new system is the I/O charge. In order to reduce excess charges, we blocked our records into 20,620 character maximum, and reduced our I/O operations accordingly. This also reduced the length of our history tape. As of July,1, 1970, costs have gone up by fifty per cent (50%).

The total costs to the end of April 1970, equals $7,445.89. We used to spend $10,000.00 annually on buying LC cards. Although this year MARC didn't supply us with all our LC cards, nevertheless, $10,000 per annum will go a long way to defraying the costs of automatic card set production.

The Target date for producing card sets is November when our automated Cataloguing/Acquisition Systems starts. The programs are already written and are presently being tested. It's too early to predict the costs of card set production, but the printing costs now are at around ten cents per unit card. Therefore, a set might run between 60¢ to 90¢, at current rates.

---

* See figure 37

Uncontrollable Costs

The fourth edition of the MARC Users Manual has just come out. There are format changes in this edition that have caused us to make changes to our programs. Although these changes have not yet accounted for any high costs to date, however, if we were to take advantage of these changes, then considerable costs could be incurred.

One change is the first step towards accommodating non roman alphabetic characters. Although the present ASCII code configuration could accommodate some additional graphics, there are too few unused positions, of the possible 256 binary representations, to provide enough codes for non roman alphabet characters such as Greek, Arabic, etc. For this reason, Greek, subscript and superscript characters have been placed in separate character sets. These separate character sets will be indicated by locking escape sequences.

All records will begin in the standard set. When an escape is made to another character set, all characters following the escape will be interpreted as being part of the variant character set until another escape sequence is reached or the end of the record is reached. Presumably the new A.L.A. print change will contain some Greek characters.

Another change as indicated in the fifth addendum is a new status code called R for replace. This status code was a new code to take care of replacing a record on the history tape, which has the same LC card number as on the new weekly tape but with different bibliographic data on the new weekly tape record. Our update program was modified to accommodate this status code but it has never been used by LC and therefore, we have come to the conclusion that the same effect is possible with LC simply putting through a new record with a C status code.

Illustration 38 indicates how backward slashes in sets of three could be used in the MARC format to indicate a print statement as separate from a sort statement. The Library of Congress have not, as of this time, adopted this method of indicating sort fields, but if they do, then our local programmes will have to be changed to accommodate this new feature.

It appears that three alternatives are available to users. (a) The user library could delete the filing statement during translation. In this case, the file and tape would be shortened and the user library would have to change the record directory. (b) One could retain the entire record but suppress the filing statement when printing catalogue cards. In this instance, no reformatting would be necessary but the backward slashes would remain there if one wanted to use them in the future. This is what we plan to do at the University of Saskatchewan. (c) The user library could set up a filing instruction and store it in the 900 fields which have been left empty for the user's own applications. If the backward slashes were put there, it would lengthen the file over alternative (a) and the user would have to change his record directory.

As presently conceived, data between the first and second backward slash, in a set of three, would be used for printing only. Data elements contained by the second and third backward slash would be used for filing purposes alone. Any data that is not contained within the set of three backward slashes, will serve both printing and sorting purposes.

How to Reduce Costs

Costs can be reduced by (a) co-operative cataloguing. The high cost of maintaining data bases[7] will lead to more cost sharing in the area of cataloguing. L.C. estimates about $45,000/month[+] for disc storage of one million bibliographic records. An example of such a cooperative effort is the transmission of MARC data, from our main data base in Saskatoon to the computer centre in Regina. Ed will have more to say about this in a few minutes. (b) Cost could be further reduced by decreasing the size of the history tape or the main file. In order to do this, one could use a time decay method, such as dropping any record which is one year or older.

Initially we tried to do a high usage study, and as originally conceived, this study was to indicate to us, if we could expect to be placing orders for, or receiving copies of books within twenty weeks after LC had created a MARC record for the item. So far this failed to be the case, and perhaps no library can structure its collection development policy so that it will have made a decision to purchase a book within twenty weeks (or some other specified time) after a book is first published. This type of study may, however, turn out to be valid after LC gets its own production schedules running more smoothly. Certainly during the first year of MARC distribution the tapes have contained records which have been in process for more than the previous week. Another way to decrease the history tape is through deselection by subject. For instance, approximately 5% of the records on MARC tapes refer to children's literature. Perhaps these records could be deleted from the local history tapes of the user library. (c) Another means of reducing costs is by having better access to the MARC tapes and therefore utilizing them to a greater extent and this is where our

---

*See RECON study

use of author/title codes has already increased total usage.
(d) Costs could be further reduced by greater utilization of the tapes through a MARC listing used as a faculty alerting service to current publications. This type of service could be prepared for individual departments or professors.

As Georg Mauerhoff mentioned yesterday, in cooperation with the National Science Library, i.e. utilizing their CAN/SDI programmes, we are looking at the possibility of developing an SDI or current awareness system based on MARC tapes. Known by the acronym SELDOM, i.e. SELective Dissemination Of MARC, the system will enable a greater use to be made of the tapes. This is in keeping with the aims of the National Library, as mentioned by Dr. Guy Sylvestre in his recent address to the Ottawa Chapter of the Canadian Association for Information Science, when he said "that we in Canada should endeavour to develop a MARC format which would provide a more sophisticated subject approach to the literature."

Project SELDOM will go beyond the traditional search programs, i.e. those retrieving on the Library of Congress Catalog Card Number, SBN and author/title compression codes, and permit searches on subject information, as well as author information. Since this information is stored in a variety of places in the MARC record, our plans call for searches to be conducted against personal author, corporate author, title, classification numbers (Dewey and LC), subject headings, and date. We envisage such a service for faculty use, since as Dr. Cuadra mentioned yesterday, librarians already have access to new publications through Publishers Weekly and other tools.

At the present time this type of service is provided only by the Library of Congress for its Legislative Reference Service (LRS) and Kenneth J. Bierman at the Oklahoma Department of Libraries, who has been operating a MARC-based SDI system. The latter unfortunately searches by classification range only.

Problems

With any evolving system, problems continue to arise. I like to think of an ongoing system as being run by the plant engineer after the research and development office have handed it over for the day to day use. As everyone knows, plant engineers have daily problems and no one really expects any system to run perfectly.

First, I would like to say a few words about outside problems, or problems which are beyond the control of the user library. I don't think anyone anticipated the first problem we ran into. The MARC tapes are produced and distributed from Argonne National Laboratory in Illinois, and in that reproduction process, the parity check on the first tape failed. This meant that on the very first tape we received from them, we were unable to read it on our tape drives. Our systems people finally discovered the error, and by sending the tape to another 360 installation in town which was also unable to read it, thereby verified the production defect. (Bill Newman, I'm sure could elaborate on this problem later, if any of you are interested)

Content changes by L.C. are beyond the control of the user library. Changes such as I previously discussed, as the inclusion of the filing statement. The size of the data bank or the number of MARC entries produced by L.C. are also beyond the control of the user library.

At this time, there is still no well defined communication pattern for handling errors, I suggested that an error detection routine be established, whereby the first user finding an error on a particular tape, would advise L.C. immediately and L.C. in turn would notify all user libraries.

L.C. has been notifying users of any changes that they are making soon enough, but the problem to the user library is to know when the change has actually gone into effect and when to expect the change on the tapes. For instance, we have never found an 'R' status record on the tapes to date.

The other type of problem is the user problem, and one which every user library must determine for itself. That is, the kind of output that is required, for example, card printouts, book catalogs, S.D.I. service, etc. All these products will require production schedules, and format decisions, which demand a lot of time and effort before a consensus is reached on the final output.

Very often it is rather difficult to discuss the pros and cons of a particular output with the user group until you have something to show them. This is basically a selling job, and the most difficult thing in the world to sell is an abstract idea. Therefore, very often considerable development costs are necessary before you can even show your organization a sample product.

Future

Last week we started transmitting MARC data to the Computer Centre on the Regina Campus and this new exploitation of our MARC data base has been enthusiastically received by the Regina Campus library.

Regular requests for MARC printouts which are initiated by Regina are expected to become routine shortly. Data transmission between the two computers at the Regina and Saskatoon campuses is being undertaken in order to maximize the use of computing power at the University of Saskatchewan. Projected costs of this communication link run around $2,500.00 per month or approximately $15.00 per hour for an 8 hour day in a 20 day month. This facility will transmit 500, 100 character lines per minute. If we allowed 12 lines per catalogue card and used 2 up forms, we should be able to transmit 80 unit cards per minute. Under the above assumptions we could transmit to Regina, 80 unit card printouts at a communications cost of 25¢. On the basis of our April MARC update and translation costs of $49.81, we could have purchased 199 minutes of communications time. This would allow us to transmit approximately 16,000 unit cards.

It would seem that the distance over which a library could economically transmit and receive from a remote data base, would increase as the size of the data base grew, since communication costs are proportional to the distance of the communication and the cost of data storage increases as the size of the data base grows. Using a single data base becomes more economically attractive if the participating libraries shared the costs of developing and maintaining multiple access points to the data.

Kilgour describes the state wide library network that is being developed in Ohio in the February issue of Datamation. He makes it quite clear that he believes that such a system is economically viable. In Canada, I understand the University of Waterloo and Brock University have been using remote job entry and data transmission very successfully for the past year. The results of the data transmission experiments in Saskatchewan will provide valuable implications for the future plans of IPCUR* and indeed, the entire country. Particularly, if as Dr. Guy Sylvestre recently said, "it should not be too long before we begin to realize a national system in which all the main libraries of the country will be linked by computer terminals to data banks." He further recommended "that the department of communications insure that the needs of the federal and other research libraries be met in any national and international communications systems which it may develop."[8] (see figure 39

The Saskatoon Computer Advisory Committee is considering an on-line terminal network for our campus in late 1972. At this time, our library system would be converted to an on-line application (see figure 40) and it is obvious that the faculty and students would like to have on-line access to the library catalogue, but we have delayed any plans for converting our shelf-list at this time, because for us, this would be contingent on the RECON project becoming operational and on the availability of inexpensive on-line hardware and software facilities. You are all aware that the RECON Working Task Force has proposed a conversion strategy through to June 1976. It is unlikely that any funding agency would support a library conversion project for data already available through the L.C. RECON Project.

In Canada we might also look forward to some guidelines and possibly a MARC service from the National Library. A machine readable NUC for Canada, as Dr. Katz from our campus emphasized in his Science Council of Canada report number 6, and similar to that proposed for the United States, would be in keeping with international cooperation as vividly described by Coward, in the December 1969 issue of the Journal of Library Automation.

---

* Interprovincial Committee on University Rationalization

Perhaps this brief look at future developments for MARC has sounded like Orwell's 1984, but I really wonder if any library considering automation can do so in isolation of these current developments and projections. And particularly in view of the recommendations passed by the National Conference on Cataloguing Standards that, "the exact content of a Canadian MARC format" be investigated by the National Library.

footnotes

1. Kilgour, Frederick G., A Regional Network - Ohio College Library Centre, Datamation, 16 (no. 2) February 1970.

2. Warheit, I.A., Design of Library Systems for Implementation with Interactive Computers, Journal of Library Automation, 3 (no. 1) March 1970.

3. Bidlack, Russel E., Typewritten Catalog Cards: A Manual of procedure and form with 125 sample cards, Ann Arbor, Campus Publishers, 1966, c 1959.

4. Payne, Charles T., Comparison of LC Proofslip and MARC Tape Arrival Dates at the University of Chicago Library, Journal of Library Automation, 3 (no. 2) June 1970 pp. 115-121.

5. Ruecking, Fredrick H., Bibliographic Retrieval from Bibliographic Input; the Hypothesis and Construction of a Test, Journal of Library Automation, 1 (no. 4) December 1968 pp. 227-238.

6. Dolby, J.L. and Forsyth, V.J., An Analysis if Cost Factors in Maintaining and Updating Card Catalogues, Journal of Library Automation, 2 (no. 4) December 1969 pp. 235.

   Dolby, J.L., Forsyth, V.J. and Resnikoff, H.L., Computerized Library Catalogs: Their Growth, Cost, and Utility, Cambridge, Mass., M.I.T., 1969.

7. "Computer Requirements for a National Bibliographic Service," in RECON Working Task Force. Conversion of Retrospective Catalog Records To Machine-Readable Form; a Study of The Feasibility of a National Bibliographic Service prepared by the RECON Working Task Force. Henriette D. Avram, Chairman, William R. Nugent, Josephine S. Pulsifer, John C. Rather, Joseph A. Rosenthal, Allen B. Veaner. Washington, 1969. pp. 69.

8. Sylvestre, Guy, Brief submitted by the National Librarian to the Senate Committee on Science Policy, Ottawa, 1969. Reproduced with the permission of the Queen's Printer for Canada from Proceedings of the Special Committee on Science Policy of the Senate of Canada. pp. 37.

Some Citations Pertinent to Author-Title Compression Codes

"Computer Requirements for a National Bibliographic Service," in RECON Working Task Force. Conversion of Retrospective Catalog Records To Machine-Readable Form; a Study of the Feasibility of a National Bibliographic Service prepared by the RECON Working Task Force. Henriette D. Avram, Chairman, William R. Nugent, Josephine S. Pulsifer, John C. Rather, Joseph A. Rosenthal, Allen B. Veaner. Washington, 1969. pp. 183-223.

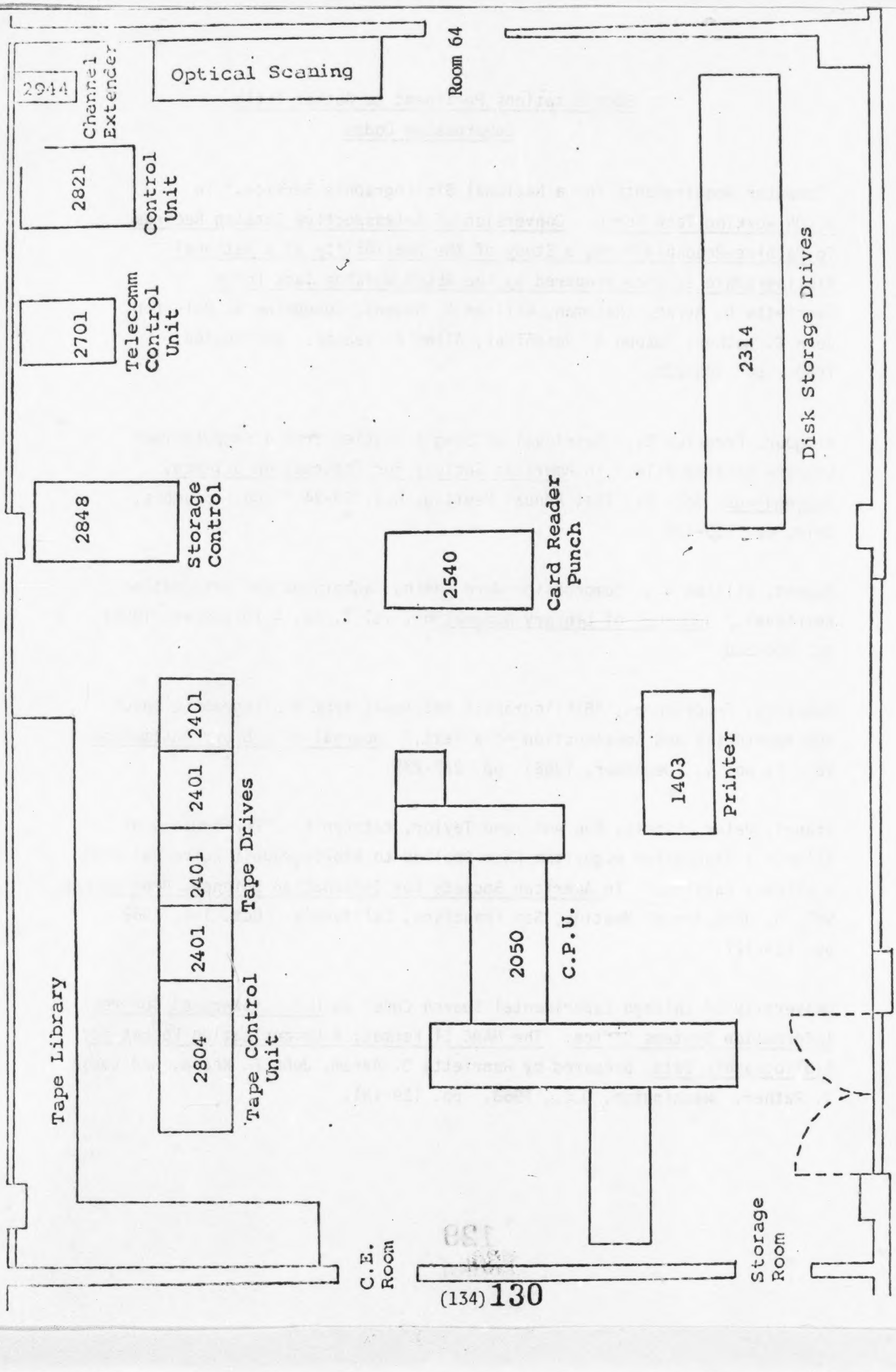
Kilgour, Fredrick G., "Retrieval of Single Entries from a Computerized Library Catalog File," in American Society for Information Science, Proceedings, Vol. 5. 31st Annual Meeting, Oct. 20-24, 1968, Columbus, Ohio, pp. 133-136.

Nugent, William R., "Compression Word Coding Techniques for Information Retrieval," Journal of Library Automation, vol 1, no. 4 (December, 1968) pp. 250-260.

Ruecking, Fredrick H., "Bibliographic Retrieval from Bibliographic Input; the Hypothesis and Construction of a Test," Journal of Library Automation. Vol. 1, no. 4, (December, 1968) pp. 227-238.

Stangl, Peter, Lipetz, Ben Ami, and Taylor, Kathryn F., "Performance of Kilgour's Truncation Algorithm When Applied to Bibliographic Retrieval from a Library Catalog," In American Society for Information Science; Proceedings, Vol. 6, 32nd Annual Meeting, San Francisco, California. Oct. 1-4, 1969, pp. 125-127.

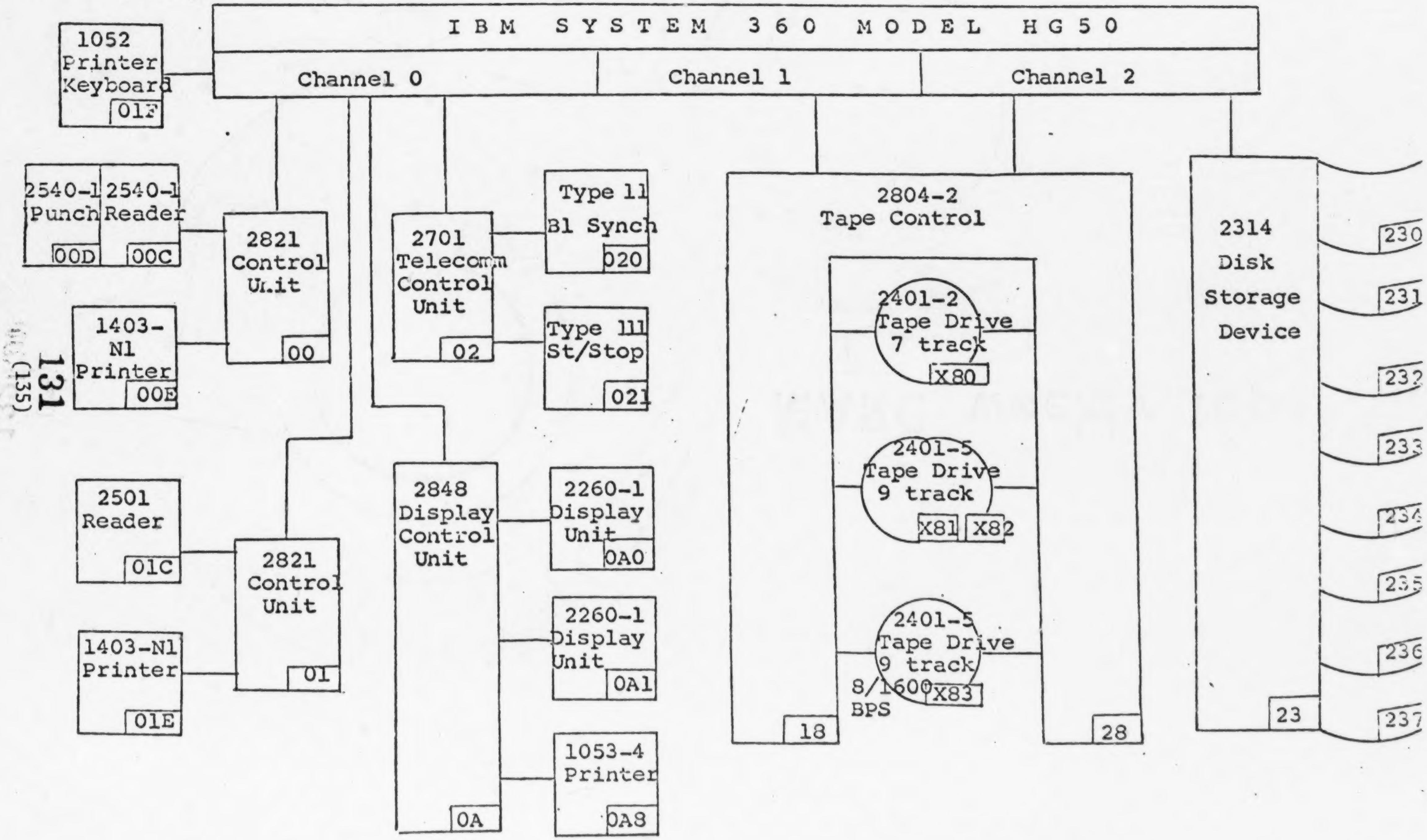
"University of Chicago Experimental Search Code" in U.S. Library of Congress. Information Systems Office. The MARC II Format; A Communication Format for Bibliographic Data, prepared by Henrietta D. Avram, John F. Knapp, and Lucia J. Rather. Washington, D.C., 1968. pp. 129-131.


Optical Scaning
Room 64
2944
Channel Extender
2821
Control Unit
2701
Telecomm Control Unit
2848
Storage Control
2540
Card Reader Punch
2314
Disk Storage Drives
1403
Printer
2050
C.P.U.
2401
2401
2401
2401
2401
Tape Drives
2804
Tape Control Unit
Tape Library
C.E. Room
Storage Room

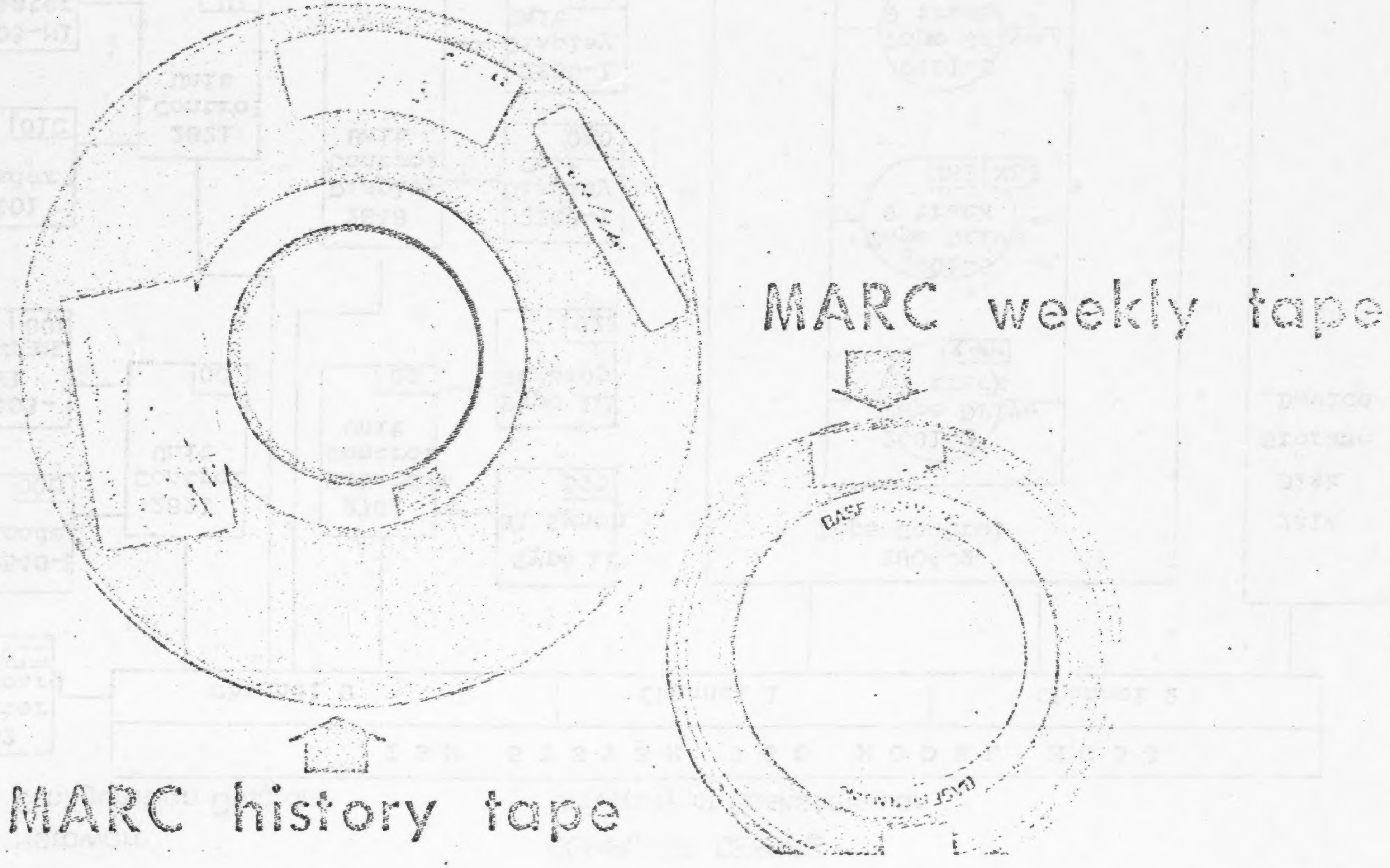
FLOOR PLAN COMPUTER ROOM
University of Saskatchewan
Room 70 Arts Addition
Scale = ¼" = 1'0"

Hardware
Configuration Diagram

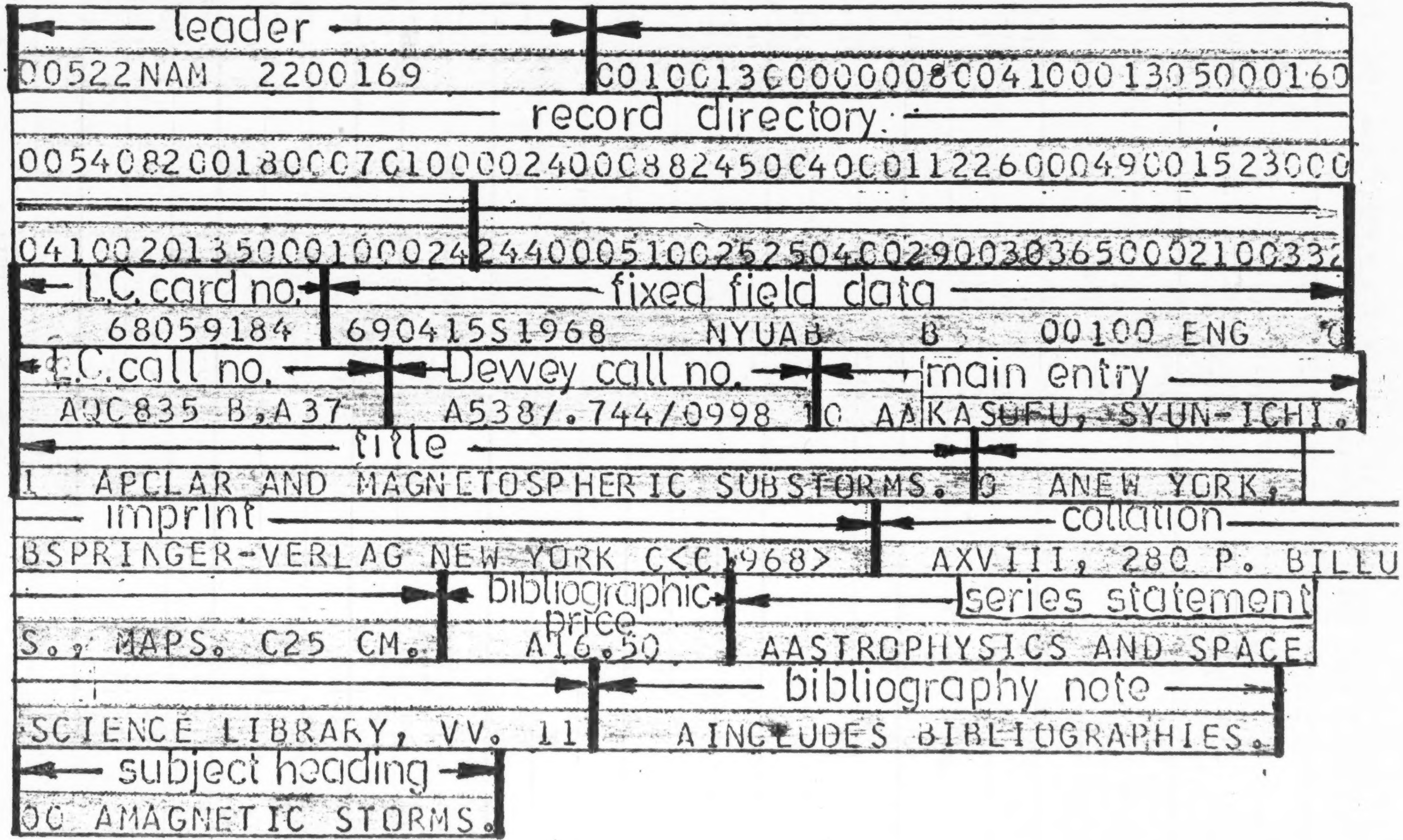
COMPUTER CENTRE
University of Saskatchewan


IBM SYSTEM 360 MODEL HG50
Channel 0
Channel 1
Channel 2
1052 Printer Keyboard
01F
2540-1 Punch
00D
2540-1 Reader
00C
2821 Control Unit
00
1403-N1 Printer
00E
2501 Reader
01C
1403-N1 Printer
01E
2821 Control Unit
01
2701 Telecomm Control Unit
02
Type 11 B1 Synch
020
Type 111 St/Stop
021
2848 Display Control Unit
0A
2260-1 Display Unit
0A0
2260-1 Display Unit
0A1
1053-4 Printer
0A8
2804-2 Tape Control
18
28
2401-2 Tape Drive 7 track
X80
2401-5 Tape Drive 9 track
X81
X82
2401-5 Tape Drive 9 track
8/1600 BPS
X83
2314 Disk Storage Device
23
230
231
232
233
234
235
236
237

Library University of Saskatchewan


MARC weekly tape
BASF
MARC history tape

```
LIBRARY OF CONGRESS * MARC-II DISTRIBUTION * MARCH 26, 1970  PAGE 2
   RECORDS THIS PAGE:   70-89135   - 71-222301

        70-89135              70-105022              70-628790
        70-89320              70-105120              70-629636
        70-89387              70-107003              70-904936
        70-90047              70-108418              70-904979
        70-90473              70-108709              70-907729
        70-90484              70-109100              70-930618
        70-90536              70-109895              71-9473
        70-90800              70-110903              71-9587
        70-91004              70-112750              71-9682
        70-91694              70-113478              71-9690

        70-92454              70-441196              71-9997
        70-92801              70-441878              71-10239
        70-92734              70-442253              71-75240
        70-92750              70-442587              71-77107
        70-92821              70-442612              71-78321
        70-92848              70-442620              71-83291
        70-92861              70-442631              71-84047
        70-93956              70-442647              71-86313
        70-93967              70-442793              71-86679
        70-94350              70-442899              71-86687

        70-94511              70-442960              71-86940
        70-94893              70-443332              71-87049
        70-95393              70-443924              71-88125
        70-95611              70-443932              71-88673
        70-95760              70-444597              71-90026
        70-96170              70-444766              71-90452
        70-96189              70-444967              71-90507
        70-96197              70-444972              71-91132
        70-96765              70-446176              71-91417
        70-97398              70-446350              71-92481

        70-97868              70-446410              71-92737
        70-97923              70-446448              71-93793
        70-98431              70-446456              71-94293
        70-98939              70-446681              71-94687
        70-98944              70-446787              71-95562
        70-99199              70-446807              71-96405
        70-99203              70-447600              71-97516
        70-99463              70-601489              71-97717
        70-100113             70-603152              71-98931
        70-100428             70-603207              71-99640

        70-101208             70-604315              71-100100
        70-101354             70-604807              71-100502
        70-101637             70-604859              71-101593
        70-102989             70-605122              71-102172
        70-103034             70-605146              71-105324
        70-103205             70-605228              71-107345
        70-103221             70-605279              71-107599
        70-103229             70-605320              71-108421
        70-104782             70-626654              71-110917
        70-104991             70-627997              71-222301


                 C = CORRECTION, D = DELETION
```


leader
record directory
L.C. card no.
fixed field data
L.C. call no.
Dewey call no.
main entry
title
imprint
collation
bibliographic price
series statement
bibliography note
subject heading

```
00522NAM  2200169        C010C013C0000008C041000130500016C
0054082001800C07C1000024000882450C40C011226000490015230CC
04100201350001000242440005100252504C029003036500021003 3
  68059184  690415S1968      NYUAB    B    00100 ENG    C
 AQC835 B,A37  A538/.744/0998 1C AAKASUFU, SYUN-ICHI.
1  APCLAR AND MAGNETOSPHERIC SUBSTORMS. 0  ANEW YORK,
BSPRINGER-VERLAG NEW YORK C<C1968>  AXVIII, 280 P. BILLU
S., MAPS. C25 CM.  A16.50  AASTROPHYSICS AND SPACE
SCIENCE LIBRARY, VV. 11  AINCLUDES BIBLIOGRAPHIES.
0C AMAGNETIC STORMS.
```

HC407
.R8H87

5+6=11

Hutira, Ervin.
The development of the national economy in the Rumanian People's Republic. ~~by, Ervin Hutira.~~ Bucharest, Meridiane Pub. House, 1963.

✓82 p. illus., fold. map. 21 cm.

Errata slip inserted.

√ e ax 21my64 stss DO NOT SET unb book rush

1. Rumania — Econ. condit. — 1945-

I. Title.

330.9498 64- 4125

1

Library of Congress

UNEDITED

Edited By:

LIBRARY OF CONGRESS
MARC Input Worksheet

Languages ☐<-X for TRANSLATION

ENG

Govt Pub 1.	Conf/Meet 2.	Festschr 3.
Index X 4.	M E in body 5.	Publisher is M E 6.
Juvenile 10.	Fiction 11.	Biography 12.
Subject is M E 13.		
Pub Key 20.	Date 1 21.	
Date 2 22.	Country 23.	
Illus Forms 24.	Repro Form 25.	
Contents Forms 26.	Bib Level 27.	

S471
.S5J8

Judd, Laurence C
Dry rice agriculture in Northern Thailand.
Ithaca, N.Y., Southeast Asia Program, Dept. of Asian Studies, Cornell University, 1964.
82 p. maps 28 cm.
(Cornell Thailand Project. Interim reports series, no. 7)
Southeast Asia Program, Dept. of Asian Studies, Cornell University. Data paper no. 52.

DO NOT SET

"Derived from a doctoral dissertation entitled 'Chao rai: dry rice farmers in northern Thailand' ... Cornell University."
Bibliography: p. 73-76.

1. Agriculture--Thailand.
2. Farm life--Thailand.
3. Rice--Thailand.
I. Title.
(Series: Cornell University. Thailand Project. Interim reports series, no. 7. Series: Cornell University. Southeast Asia Program. Data paper no. 52)

630.1 64-3579

Library of Congress

Edited by: Sim

LIBRARY OF CONGRESS
MARC Input Worksheet

EDITED

Languages LAN ☐ ← X for TRANSLATION

ENG

FFD

Govt Pub 1.	Conf/Meet 2.	Festschr 3.
Index 4. X	M E in body 5.	Publisher is M E 6.
Juvenile 10.	Fiction 11.	Biography 12.
Subject is M E 13.		
Pub Key 20. S	Date 1 21. 1964	
Date 2 22.	Country 23. NYU	
Illus Forms 24.	Repro Form 25.	
Contents Forms 26. A	Bib Level 27.	

CAL	S471 .S5ǂJ8
MEPS	Judd, Laurence C.
TILA	Dry rice agriculture in Northern Thailand.
IMP	Ithaca, N.Y.,ǂ Southeast Asia Program, Dept. of Asian Studies, Cornell University,ǂ 1964.
COL	82 p.ǂ maps ǂ28 cm.
SERD	(Cornell Thailand Project. Interim reports series, No. 7)
SERD	Southeast Asia Program, Dept. of Asian Studies, Cornell University. Data paper no. 52.
	DO NOT SET
BHN	"Derived from a doctoral dissertation entitled 'Chao rai: dry rice farmers in northern Thailand' ... Cornell University."
BIB	Bibliography: p. 73-76.
SUTǂZ	1. Agriculture--Thailand.
SUTǂZ	2. Farm life--Thailand.
SUTǂZ	3. Rice--Thailand. ~~I. Title.~~ e
SACN	e ~~(Series:~~ Cornell University.ǂ Thailand Project.ǂ Interim reports series,ǂ no. 7.
SACN	e ~~Series:~~ Cornell University.ǂ Southeast Asia Program.ǂ Data paper no. 52)
	DDC 630.1
	CRD 64-3579
	Library of Congress

LAN	eng
FFD	20.s 21.1964 23.nyu 26.a 4.x
cal	S471.S5;J8
meps	Judd, Laurence C.
tila	Dry rice agriculture in no Northern Thailand.
imp	Ithaca, N.Y.,;Southeast Asia Program, Dept. of Asian Studies, Cornell University,;1964.
col	82 p.;maps;28cm.
serd	Southeast Asia Program, Dept. of Asian Studies, Cornel Un 1 University. Data paper no. 52.
bhn	"Derived from a doctoral dissertation entitled 'Chao rai: dry rice farmers in northern Thailand' ... Cornell University."
bib	Bibliography: p. 73-76.
sut;z	Agriculture--Thailand.
sut;z	Farm life--Thailand.
sut;z	Roce--Thailand (*This information will be deleted and replaced by its corrected form*) Rice--Thailand.
sacn	Cornell University.;Thailand Project.:Interim reports series,;no. 7.
sacn	Cornell University.;Southeast Asia Program.:Data paper;no. 52)
ddc	630.1
crd	64-3579

```
CRD/1      #a#64-003579
FFD        1.        2.        3.       4.x       5.        6.
           10.       11.       12.      13.       20.s      21.1964
           22.       23.nyu    24.b     25.       26.       27.
           28.
LAN/2      #a#eng
CAL/3      #ab#S471.S5#J8
DDC/4      630.1
MEP/1      1#a#Judd, Laurence C.
TIL/1      1#a#Dry rice agriculture in Northern Thailand.
IMP/2      #abc#Ithaca, N.Y.,#Southeast Asia Program, Dept. of Asian Studies
           , Cornell University,#1964.
COL/1      #abc#82.p.#maps#28 cm.
SER/1      2#a#Southeast Asia Program, Dept. of Asian Studies, Cornell
           University. Data paper no. 52.
BHN/1      #a#"Derived from a doctoral dissertation entitled 'Chao rai:  dry
           rice farmers in northern Thailand' ... Cornell University."
BIB/2      #a#Bibliography:  p. 73-76.
SUT/1      -0#az#Agriculture--Thailand.
SUT/2      -0#az#Farm life--Thailand.
SUT/3      -0#az#Rice--Thailand.
SAC/1      2#abtv#Cornell University.#Thailand Project.#Interim reports series
           ,#no. 7.
SAC/2      2#abtv#Cornell University.#Southeast Asia Program.#Data paper#no.
           52
```

L.C. Tagging Scheme (A Sample of Some of the Tags and Indicators that are Used)

Tag & Indicators (2)	Variable Field Data Element
	Knowledge Numbers
050 0 ƀ	LC Call Number - book is in LC
1 ƀ	book is not in LC
051 ƀ ƀ	Copy Statement
*060	NLM Call Number
082 ƀ ƀ	Dewey Decimal Classification number
	Main Entry
100 00	Personal Name - Forename - Main entry is not subject
01	- Main entry is subject
10	- Surname - Main entry is not subject
11	- Main entry is subject
20	- Multiple Surname - Main entry is not subject
21	- Main entry is subject
30	- Name of Family - Main entry is not subject
31	- Main entry is subject
110 00	Corporate Name - Surname (inverted) - Main entry is not subject
01	- Main entry is subject
10	- Place or Place & Name - Main entry is not subject
11	- Main entry is subject
20	- Name (direct order) - Main entry is not subject
21	- Main entry is subject
111	Corporate Name - Conference or meeting Note - Field 111 for conference or meeting headings is a subdivision of field 110. Therefore, the indicators used in this field are the same as those used for any corporate name.
130 ƀ0	Uniform Title Heading - Main entry is not subject
ƀ1	- Main entry is subject

---

* The Library of Congress will not supply data for these fields at present.

A SAMPLE OF MARC DELIMITER CODES

050 - LC, NLM, and NAL Call Numbers
060
070

$a - Class Number
$b - Book Number

051 LC Copy Statement

$a - Class Number
$b - Book Number
$c - Copy Information

100 - Personal Name
400
600
700
800

$a - Name
$b - Numeration
$c - Titles and other words associated with the name
$d - Dates
$e - Relator
$k - Form Sub-heading
$t - Title (of book)

110 - Corporate Name
410
610
710
810

$a - Name
$b - Each subordinate unit in hierarchy
$e - Relator
$k - Form Sub-heading
$t - Title (of book)

111 - Conference or Meeting
411
611
711
811

$a - Name
$b - Number
$c - Place
$d - Date
$e - Subordinate unit in Name
$g - Other miscellaneous information
$k - Form Sub-heading
$t - Title (of book)

RELATION OF TAGGING SCHEME TO FILING

Name	Numeration	Title
#aAlexander	#bIII,	#cKing of Albania
#aAlexander	#bII,	#cKing of Bulgaria
#aAlexander	#bI,	#cKing of Russia

delimiter codes

PERSONAL NAMES AS AUTHOR AND SUBJECT IN SEPARATE CATALOGUES

Tag			
100	Main Entry	Smith, John	author/title catalog
700	Added Entry	Smith, John	
600	Subject Entry	Smith, John	subject catalog

```
                                        UNIVERSITY OF SASKATCHEWAN

                                   MURRAY MEMORIAL LIBRARY - SELECTED MARC RECORDS
                                   ***********************************************
      LIBRARY OF CONGRESS CARD NUMBER    68016482
      *******************************

      CONTROL AND KNOWLEDGE NUMBERS
      *****************************

LC CALL NUMBER: DD261 .H26

      MAIN ENTRY
      **********

PERSONAL NAME: HANHARDT, ARTHUR MONROE, 1932-

      SUPPLIED TITLES AND TITLE PARAGRAPH
      ***********************************

TITLE: (TITLE IS ADDED ENTRY)
      THE GERMAN DEMOCRATIC REPUBLIC <BY> ARTHUR M. HANHARDT, JR.

IMPRINT: BALTIMORE, JOHNS HOPKINS PRESS <1968>

      COLLATION
      *********

COLLATION: XXI, 126 P. MAP. 21 CM.

BIBLIOGRAPHIC PRICE: 6.00

      SERIES NOTES
      ************

TITLE (TRACED SAME): INTEGRATION AND COMMUNITY BUILDING IN EASTERN EUROPE

      BIBLIOGRAPHIC NOTES
      *******************

GENERAL NOTES: "JH-EEL."

BIBLIOGRAPHY NOTE: BIBLIOGRAPHY: P. 125-126.

      SUBJECT ADDED ENTRIES AND LC SUBJECT HEADINGS
      *********************************************

POLITICAL JURISDICTIONS ALONE OR WITH SUBJECT SUBDIVISIONS: GERMANY (DEMOCRATIC REPUBLIC, 1949- )
```

First Style
May 1, 1969

```
                              UNIVERSITY OF SASKATCHEWAN

                  MURRAY MEMORIAL LIBRARY - ACQUISITIONS MARC RECORDS
                  ****************************************************

LC CARD NUMBER:    68009755
LC CALL NUMBER: HD8514 .W54

MAIN ENTRY
PERSONAL NAME: WINDMULLER, JOHN P.

TITLE: (TITLE IS ADDED ENTRY)
     LABOR RELATIONS IN THE NETHERLANDS, BY JOHN P. WINDMULLER.
IMPRINT: ITHACA, N.Y., CORNELL UNIVERSITY PRESS <1969>

COLLATION: XVI, 469 P. 24 CM.
BIBLIOGRAPHIC PRICE: 14.00

BIBLIOGRAPHY NOTE: BIBLIOGRAPHY: P. 442-464.

SUBJECT ADDED ENTRIES AND LC SUBJECT HEADINGS
TOPICAL: INDUSTRIAL RELATIONS NETHERLANDS HISTORY.
```

Second Style

May 15, 1969

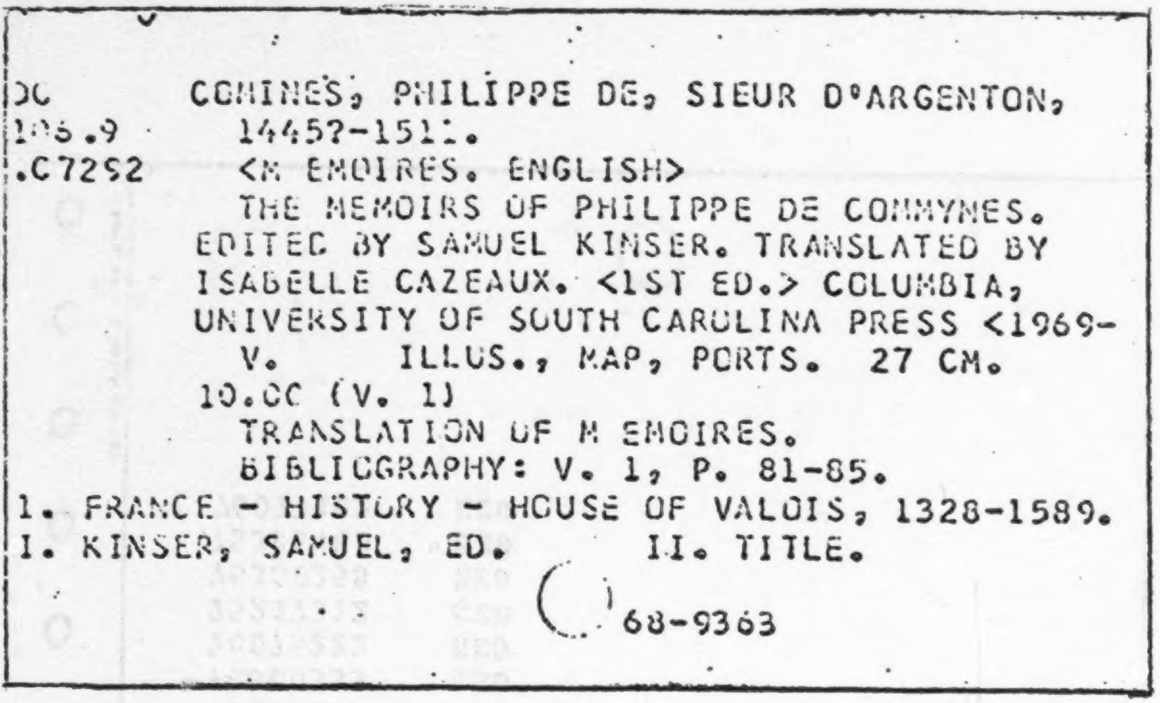
```
.................................................
.                                               .
.                                               .
.Z       JENNINGS, JOHN MELVILLE.               .
.733        THE LIBRARY OF THE COLLEGE OF WILLIAM .
..W686J4 AND MARY IN VIRGINIA, 1693-1793 <BY> JOHN.
.        M. JENNINGS. CHARLOTTESVILLE, UNIVERSITY .
.        PRESS OF VIRGINIA <1968>               .
.           IX, 91 P.  ILLUS., FACSIMS., PORTS.  25.
.        CM.  (LIBRARY CONTRIBUTIONS, NO. 6)    .
.           "PUBLISHED FOR THE EARL GREGG SWEM  .
.        LIBRARY OF THE COLLEGE OF WILLIAM AND  .
.        MARY IN VIRGINIA."                     .
.           INCLUDES BIBLIOGRAPHICAL REFERENCES. .
.1. WILLIAM AND MARY CO      LLEGE, WILLIAMSBURG, .
.VA. EARL GREGG SWEM LI      BRARY.  I. TITLE.  .
. (SERIES: WILLIAM AND       MARY COLLEGE,      .
.                        (CONTINUED ON NEXT CARD) .
.................................................
```

```
.................................................
.                                               .
.                                               .
.Z        JENNINGS, JOHN MELVILLE.  THE LIBRARY OF .
.733        THE COLLEGE OF WILLIAM AND MARY IN  .
..W686J4    VIRGINIA, 1693-1793 ... (CARD 2)    .
.                                               .
.WILLIAMSBURG, VA. LIBRARY CONTRIBUTIONS, NO. 6) .
.                                               .
.                                               .
.                                               .
.                                               .
.                                               .
.                                               .
.                                               .
.                                               .
.                          68-59130             .
.................................................
```

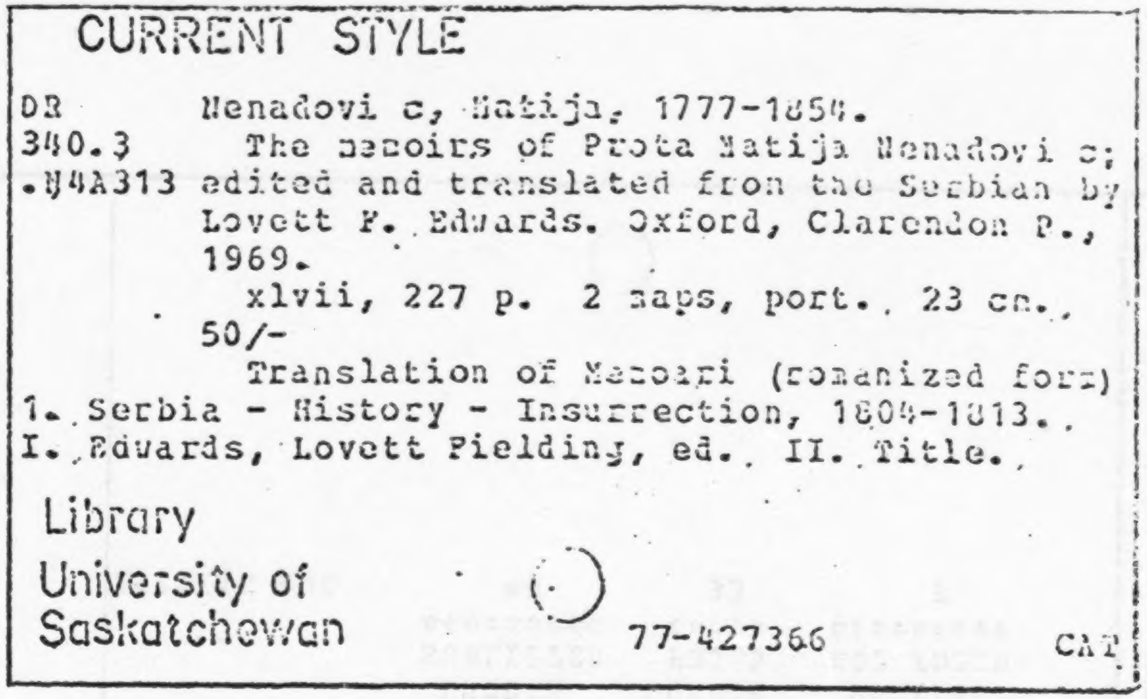
Third Style

July 1, 1969

Library University of Saskatchewan



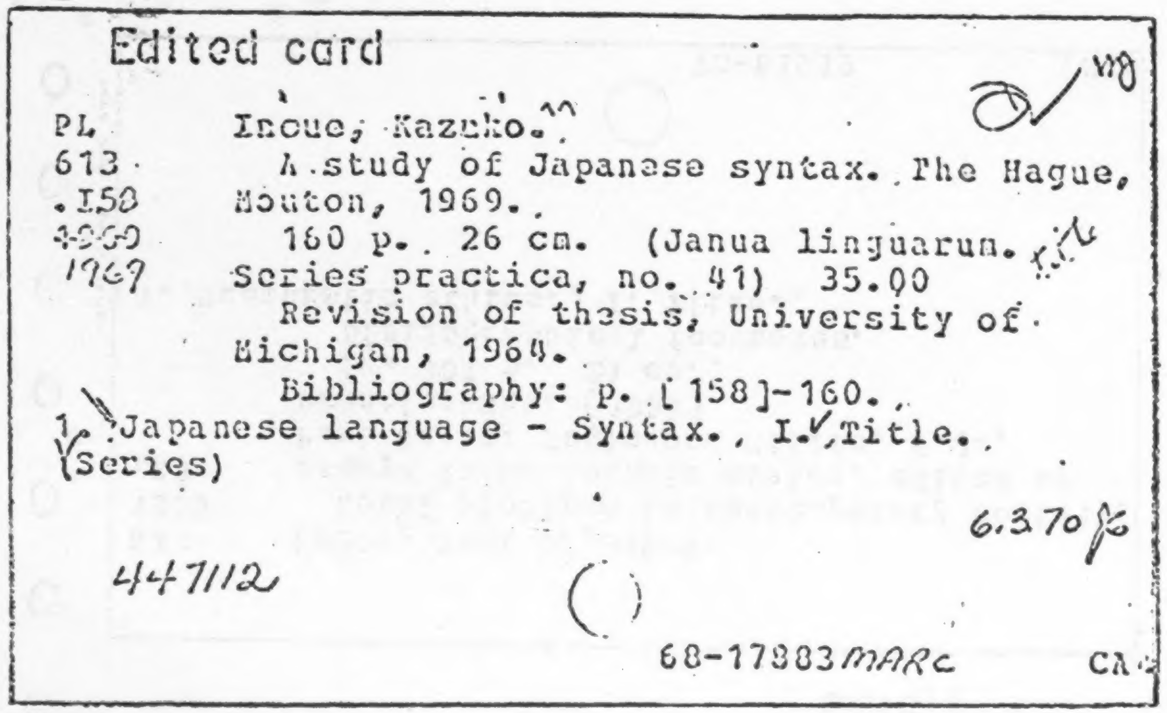
DC COMINES, PHILIPPE DE, SIEUR D'ARGENTON,
106.9 1445?-1511.
.C7292 <M EMOIRES. ENGLISH>
THE MEMOIRS OF PHILIPPE DE COMMYNES.
EDITED BY SAMUEL KINSER. TRANSLATED BY
ISABELLE CAZEAUX. <1ST ED.> COLUMBIA,
UNIVERSITY OF SOUTH CAROLINA PRESS <1969-
V. ILLUS., MAP, PORTS. 27 CM.
10.00 (V. 1)
TRANSLATION OF M EMOIRES.
BIBLIOGRAPHY: V. 1, P. 81-85.
1. FRANCE - HISTORY - HOUSE OF VALOIS, 1328-1589.
I. KINSER, SAMUEL, ED. II. TITLE.
68-9363



CURRENT STYLE

DR Nenadovi c, Matija, 1777-1854.
340.3 The memoirs of Prota Matija Nenadovi c;
.N4A313 edited and translated from the Serbian by
Lovett F. Edwards. Oxford, Clarendon P.,
1969.
xlvii, 227 p. 2 maps, port. 23 cm.
50/-
Translation of Memoari (romanized form)
1. Serbia - History - Insurrection, 1804-1813.
I. Edwards, Lovett Fielding, ed. II. Title.

Library
University of
Saskatchewan
77-427366 CAT



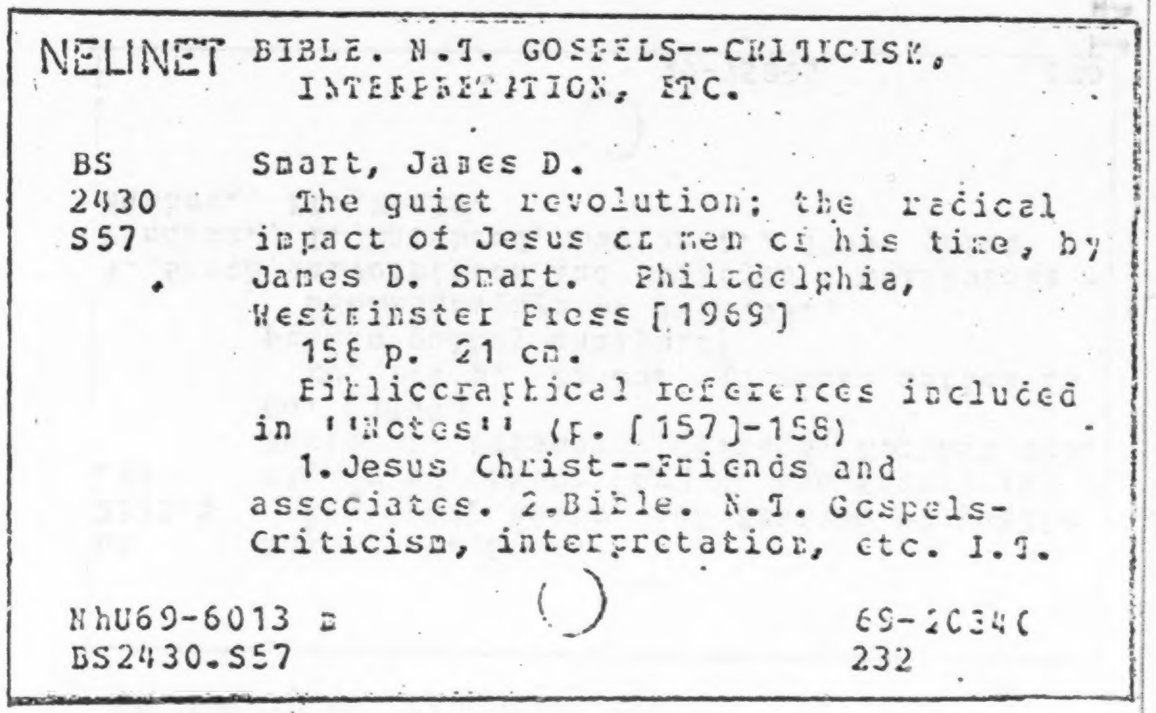
Edited card

PL Inoue, Kazuko.
613 A study of Japanese syntax. The Hague,
.I58 Mouton, 1969.
~~1969~~ 160 p. 26 cm. (Janua linguarum.
1969 Series practica, no. 41) 35.00
Revision of thesis, University of
Michigan, 1964.
Bibliography: p. [158]-160.
1. Japanese language - Syntax. I. Title.
(Series)

6.3.70

447112

68-17883 MARC CA



NELINET BIBLE. N.T. GOSPELS--CRITICISM,
INTERPRETATION, ETC.

BS Smart, James D.
2430 The quiet revolution; the radical
S57 impact of Jesus on men of his time, by
James D. Smart. Philadelphia,
Westminster Press [1969]
156 p. 21 cm.
Bibliographical references included
in ''Notes'' (p. [157]-158)
1. Jesus Christ--Friends and
associates. 2. Bible. N.T. Gospels-
Criticism, interpretation, etc. I. T.

NhU69-6013
BS2430.S57
69-20340
232

Library University of Saskatchewan

```
BJ       Kurtz, Paul W., 1925-
1360        Moral problems in contemporary society;
.K8      essays in humanistic ethics, edited by
         Paul Kurtz. Englewood Cliffs, N.J.,
         Prentice-Hall [1969]
            ix, 301 p. 21 cm.
            Bibliographical footnotes.
 1. Humanistic ethics. I. Title.

                                78-81919          MED
```

PATD BY R. L. CRAIN LIMITED

```
LB       Hansen, W. Lee.
2329.5      Benefits, costs, and finance of public
.H3      higher education [by] W. Lee Hansen and
         Burton A. Weisbrod. Chicago, Markham Pub.
         Co. [1969]
            xv, 114 p. 23 cm. (Markham series in
         public policy analysis)
            Bibliography: p. 107-109.
 1. State universities and colleges - California -
 Finance. I. Weisbrod, Burton A., 1931- joint
 author. II. Title.

                                79-75896          MED
```

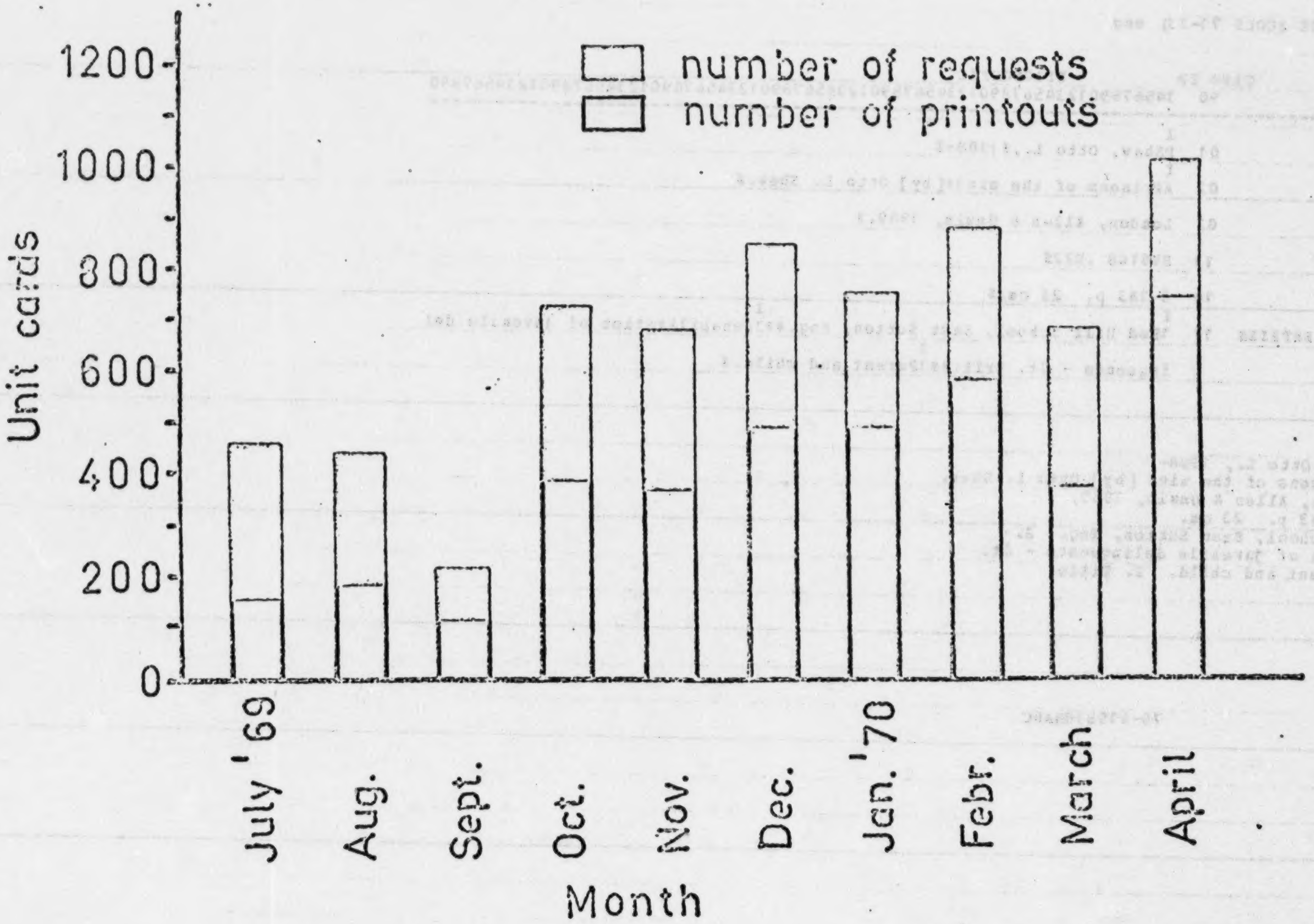
```
NO MARC RECORD FOR FOLLOWING LC CARD NUMBER(S):
    69019880      MED
    73086335      MED
    74075222      MED
    75217312      MED
    75360258      MED
    75396448      MED
    76077482      MED
```

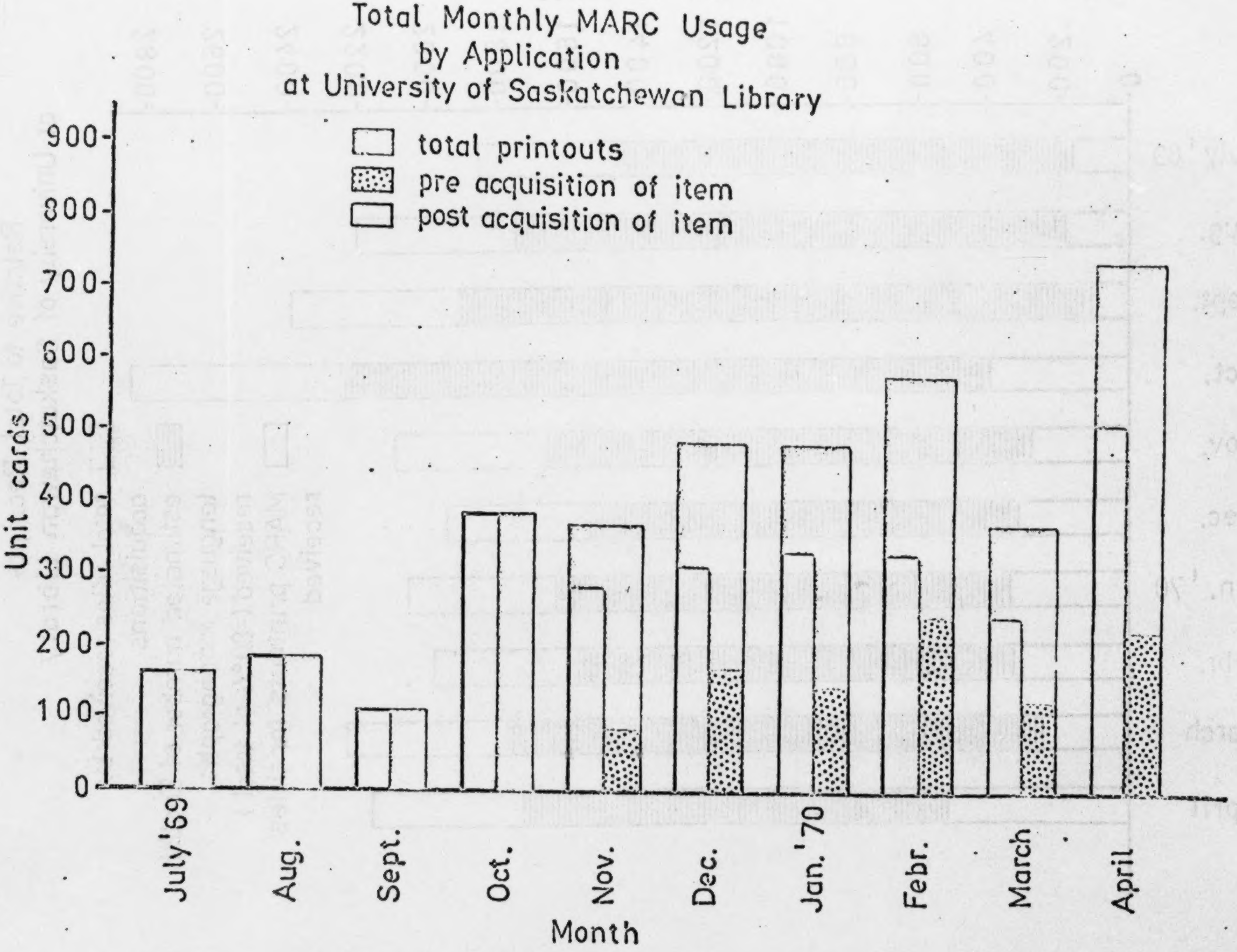
PATD BY R. L. CRAIN LIMITED

```
          MARC STATISTICS  70/02/04 MED

                      NUMBER      NUMBER    NUMBER
                      SUBMITTED   FOUND     NOT FOUND
                      *********   *****     *********
INITIAL RUN              40         33         7
```

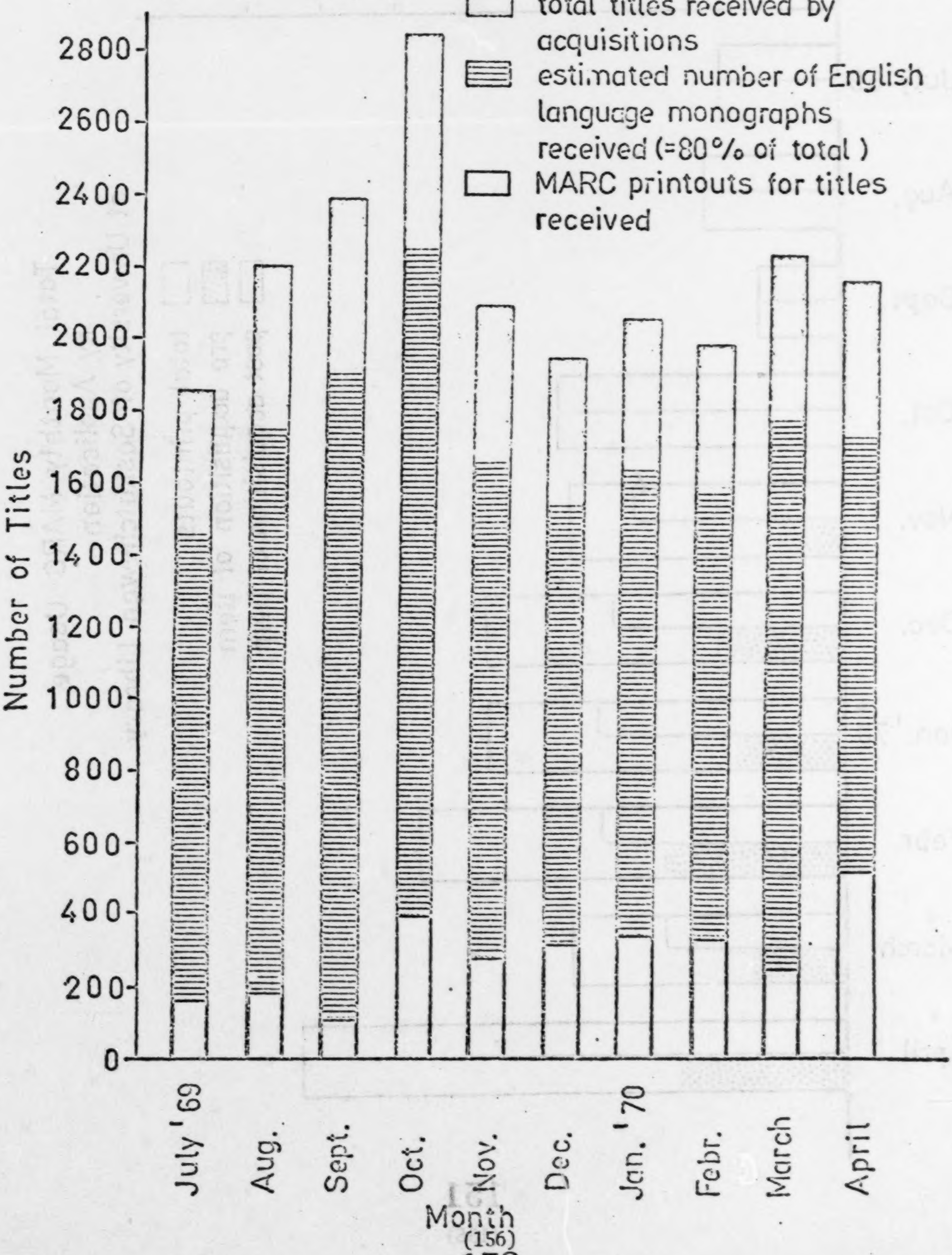
```
                          MURRAY MEMORIAL LIBRARY - CATALOGER'S LISTING
                          ITEM NO. 100081K0    DEPARTMENT NO. 038    PAGE 1

   "A" CARD FIELDS
   ---------------

AA-NO.OF COPIES (COLS 13-14) 01 AB-SBN (COLS 15-24) 041570057  AC-LC CARD NO. (COLS 25-36)    79419818  AP-ON MARC TAPE (COL 44) N

AH-LOCATION CODE (COL 50) C AK-FORM OF MATERIAL (COL 55) B AL-PRIORITY (COL 56)   AM-HANDLING CODE (COL 57)

AN-CONFERENCE OR MEETING (COL 58) N AO-NO. OF VOLUMES  (COLS 59-61) 001 AQ-COUNTRY CODE (COLS 69-70) uk

AR-LANGUAGE CODE (COLS 71-73) eng

      FIELD           CARD ID       DIAGNOSTIC
                         90  3456789012345678901234567890123456789012345678901234567890
-------------------------------------------------------------------------------------
                             I
MAIN ENTRY               01  PShaw, Otto L.,#1908-$
                             I
TITLE                    02  APrisons of the mind#[by] Otto L. Shaw.$

IMPRINT                  03  London, Allen & Unwin, 1969.$

CALL NUMBER              13  HV9148 .E22$

COLLATION                14  3-243 p.  23 cm.$
                             I                                  I
SUBJECT ADDED ENTRIES    17  1Red Hill School, East Sutton, Eng.##2Rehabilitation of juvenile del
                                                 I
                             inquents - Gt. Brit.##3Parent and child.$


HV      Shaw, Otto L., 1908-
9148       Prisons of the mind [by] Otto L. Shaw.
E22     London, Allen & Unwin, 1969.
           3-243 p.  23 cm.
1. Red Hill School, East Sutton, Eng.  2.
Rehabilitation of juvenile delinquents - Gt.
Brit.  3. Parent and child.  I. Title.


                             79-419818MARC
```

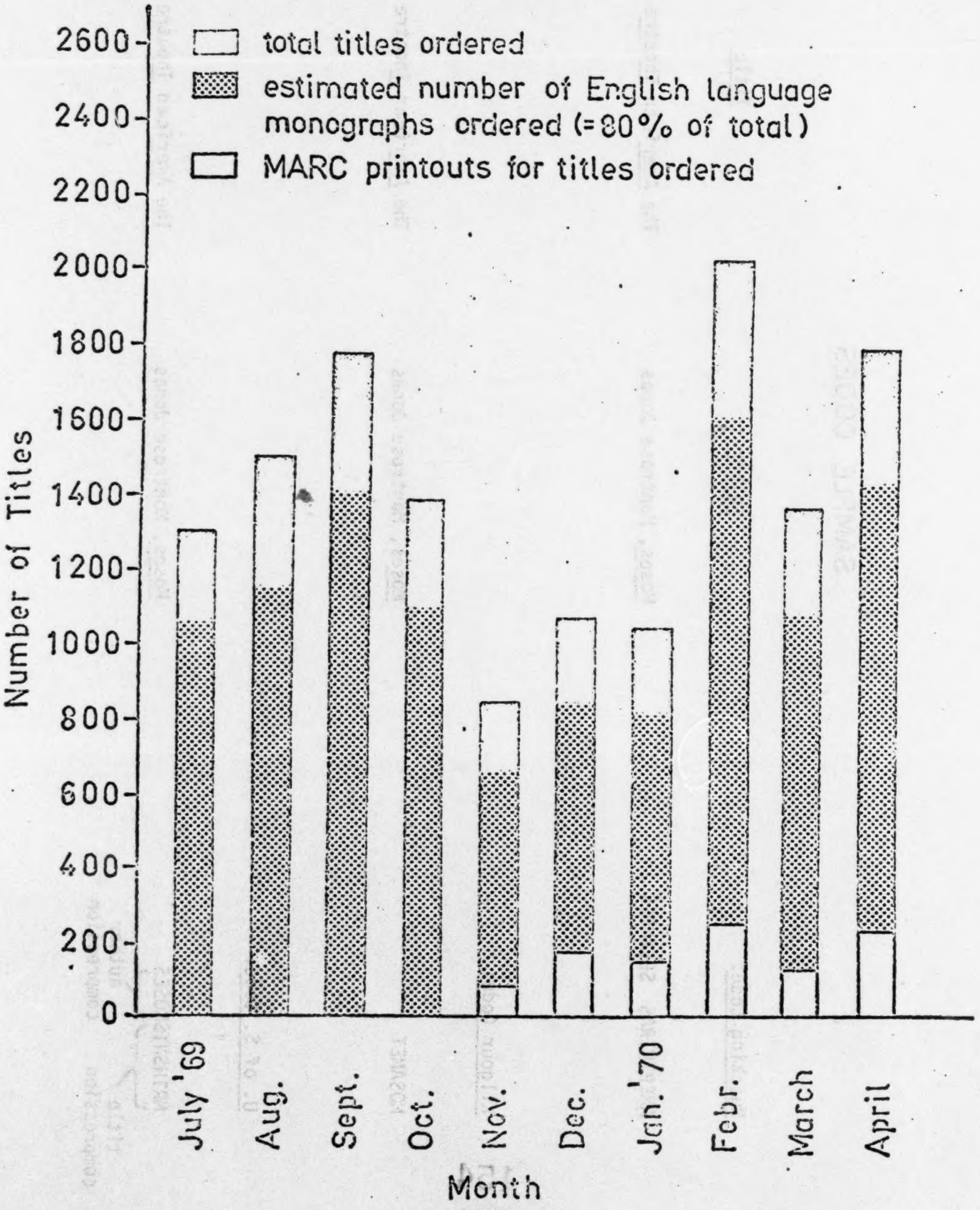
Total Monthly MARC Usage
at University of Saskatchewan Library




Total Monthly MARC Usage
by Application
at University of Saskatchewan Library
total printouts
pre acquisition of item
post acquisition of item
Unit cards
900-
800-
700-
600-
500-
400-
300-
200-
100-
0
July '69
Aug.
Sept.
Oct.
Nov.
Dec.
Jan. '70
Febr.
March
April
Month

Monthly Post Acquisitions Use of MARC
Relative to Total Receipts
at University of Saskatchewan Library



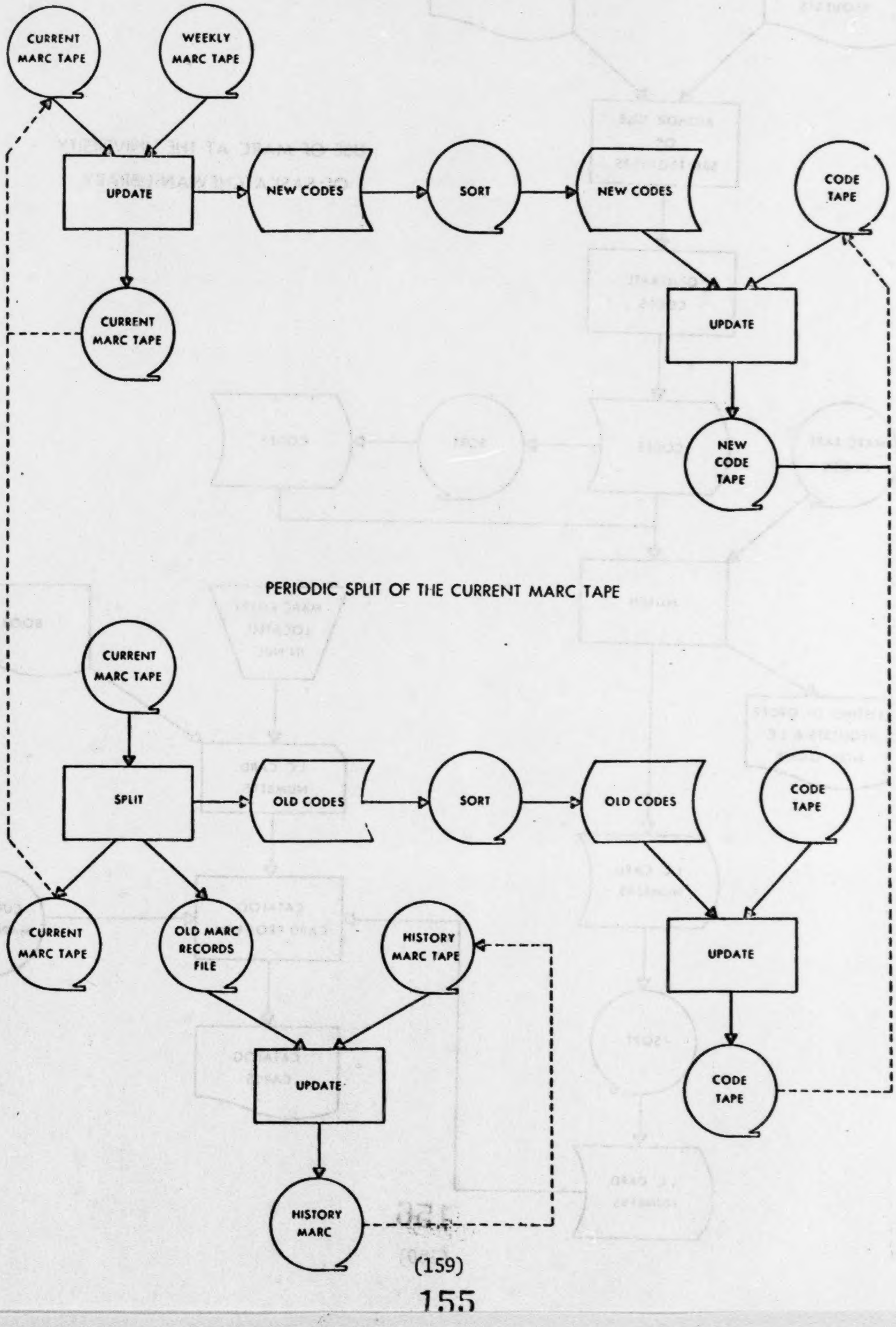


Monthly Pre Acquisitions Use of MARC Relative to Orders Placed at University of Saskatchewan Library

SAMPLE CODES

Ruecking Code:	Author	Title
AMERTHTRAS SEENMOSS	Moses, Montrose Jonas	The American Theatre As Seen By Its
Kilgour Code:		
MOSAMET	Moses, Montrose Jonas	The American Theatre As Seen By Its
U. of S. Code:		
$\underbrace{\text{MRTHSNTS}}_{\text{title compression}}\underbrace{\text{MOSES}}_{\text{author compression}}$	Moses, Montrose Jonas	The American Theatre As Seen By Its

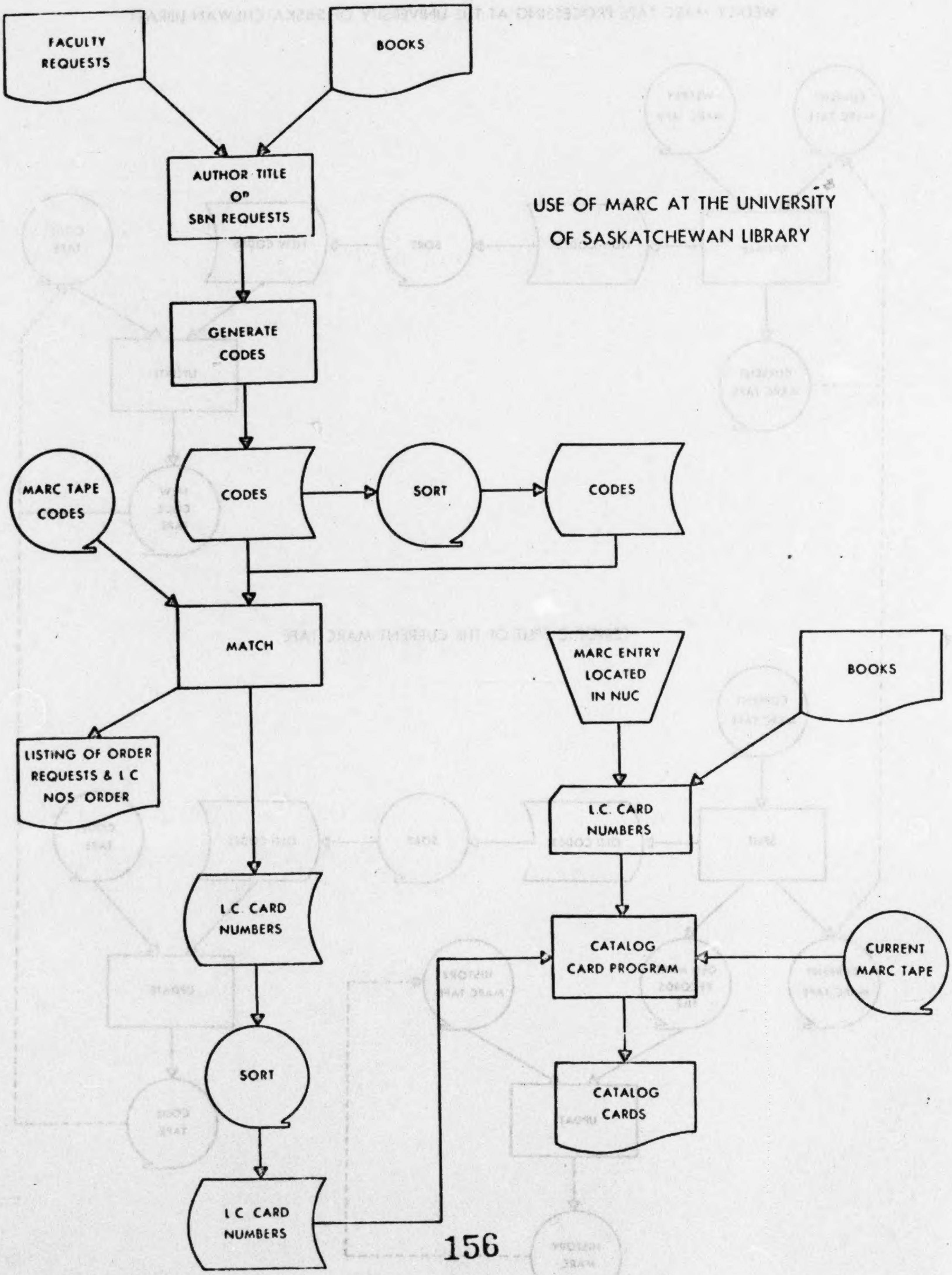
WEEKLY MARC TAPE PROCESSING AT THE UNIVERSITY OF SASKATCHEWAN LIBRARY



PERIODIC SPLIT OF THE CURRENT MARC TAPE

CURRENT MARC TAPE
WEEKLY MARC TAPE
UPDATE
NEW CODES
SORT
NEW CODES
CODE TAPE
CURRENT MARC TAPE
UPDATE
NEW CODE TAPE
CURRENT MARC TAPE
SPLIT
OLD CODES
SORT
OLD CODES
CODE TAPE
CURRENT MARC TAPE
OLD MARC RECORDS FILE
HISTORY MARC TAPE
UPDATE
UPDATE
CODE TAPE
HISTORY MARC

USE OF MARC AT THE UNIVERSITY OF SASKATCHEWAN LIBRARY

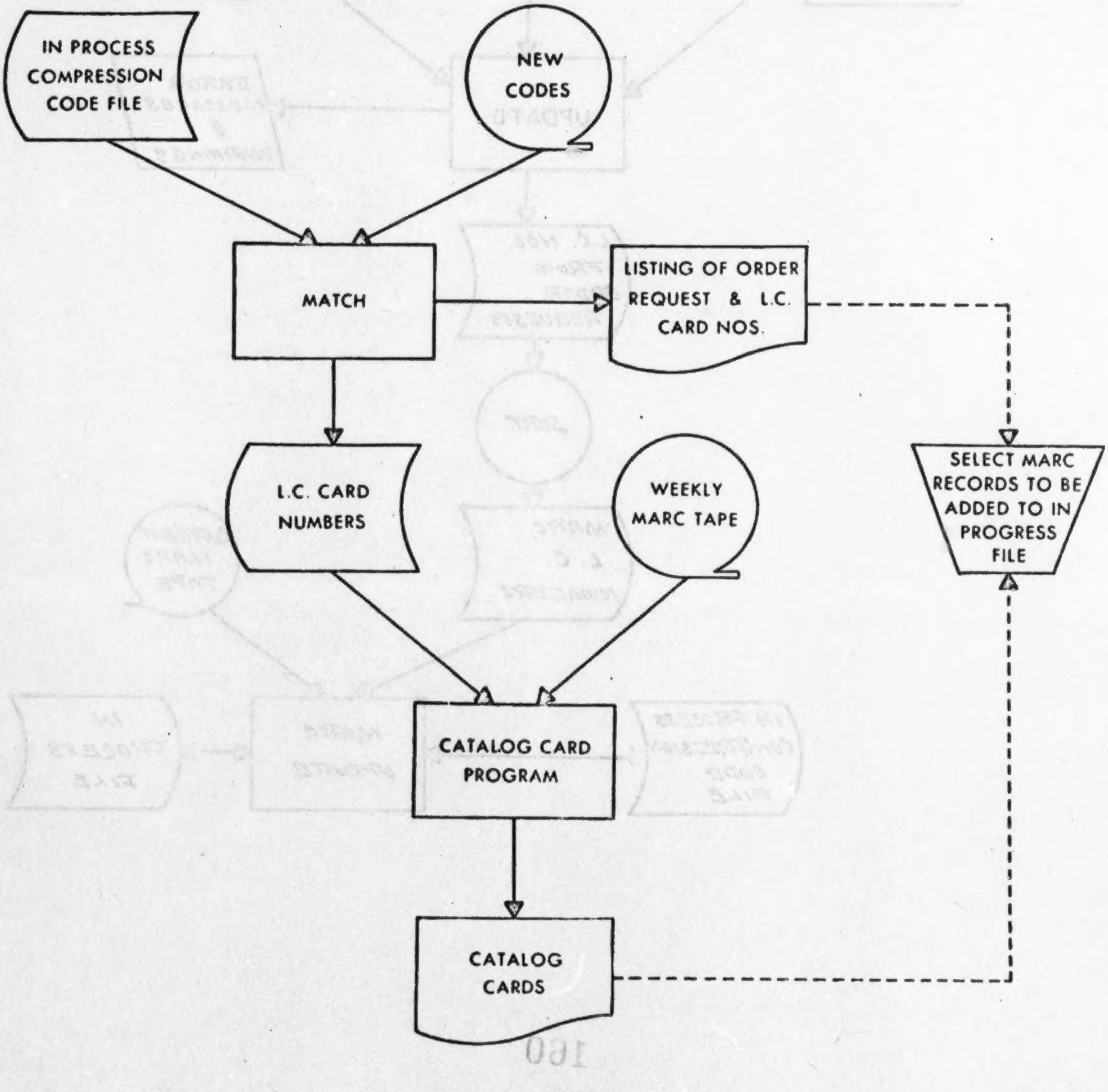

FACULTY REQUESTS
BOOKS
AUTHOR TITLE Or SBN REQUESTS
GENERATE CODES
MARC TAPE CODES
CODES
SORT
CODES
MATCH
LISTING OF ORDER REQUESTS & L C NOS ORDER
L.C. CARD NUMBERS
SORT
L C CARD NUMBERS
MARC ENTRY LOCATED IN NUC
BOOKS
L.C. CARD NUMBERS
CATALOG CARD PROGRAM
CURRENT MARC TAPE
CATALOG CARDS

```
                     LIBRARY - ENTRY/TITLE REQUESTS FOR MARC CATALOG CARDS     PAGE  1
         REQUEST                                                  SEARCH CODE          LC CARD NUMBER(S)
         -------                                                  -----------          -----------------

   HIRAI                                                       BCRGDSPLHIRAI    NO LC CARD NUMBERS FOUND
   BIOCHEMICAL REGULATION IN DISEASED PLANTS

   MACQUILLY                                                   BDDHA   MACQLY   NO LC CARD NUMBERS FOUND
   BUDDHA

   NATHAN                                                      BGSPRINGNATHAN      75081815
   THE BEGINNING SPRING

   BALDWIN                                                     BLGRXPERBALDWN   NO LC CARD NUMBERS FOUND
   THE BALDWINS AND THE GREAT EXPERIMENT

   RATTRAY                                                     BLTMBMB RATTRY   NO LC CARD NUMBERS FOUND
   THE BIOLOGICAL TIME BOMB

   MARTIN                                                      BLXSMNINMARTIN      69016344
   BELIEF EXISTENCE AND MEANING

   PINTO                                                       BNBLBTGNPINTO       71094400
   THE BIONOMICS OF BLISTER BEETLES OF THE GENES

   ROBINSON                                                    BNCLCLSTROBISN   NO LC CARD NUMBERS FOUND
   BENEFICED CLERGY IN CLEVELAND AND THE EAST RIDING

   RYDER                                                       BNKP1418RYDER       68054523
   BENIN AND THE EUROPEANS 1485-1897

   PEACOCK                                                     BRCH1818PEACCK   NO LC CARD NUMBERS FOUND
   BRADFORD CHARTISM 1838 1840

   BARLOW                                                      BS3LRDRGBARLOW   NO LC CARD NUMBERS FOUND
   OBSERVATIONS ON THE BIOLOGY OF RODENTS IN URUGUAY

   NAGERA                                                      BSPSCNLBNAGERA      74460085
   BASIC PSYCHOANALYTIC CONCEPTS ON THE LIBIDO

   RANSFORD                                                    BTSPKP  RANSRD      76444100
   BATTLE OF SPION KOP

   CAIBNALL                                                    CCHSRDVTCAIBLL   NO LC CARD NUMBERS FOUND
   ECCLESIASTICAL HISTORY OF ORDERIC VITALIS

   LUARD                                                       CHRCUMINLUARD    NO LC CARD NUMBERS FOUND
   THE CHALLENGE OF BECOMING

   BABAKHANIAN                                                 CHFNGRGUBABANN   NO LC CARD NUMBERS FOUND
   COHOMOLOGY OF FINITE GROUPS

                                                               CNCHSTOR         NO LC CARD NUMBERS FOUND
   CONTEMPORARY CHINESE STORIES

   WANG                                                        CNCHSTORWANG     NO LC CARD NUMBERS FOUND
   CONTEMPORARY CHINESE STORIES
```

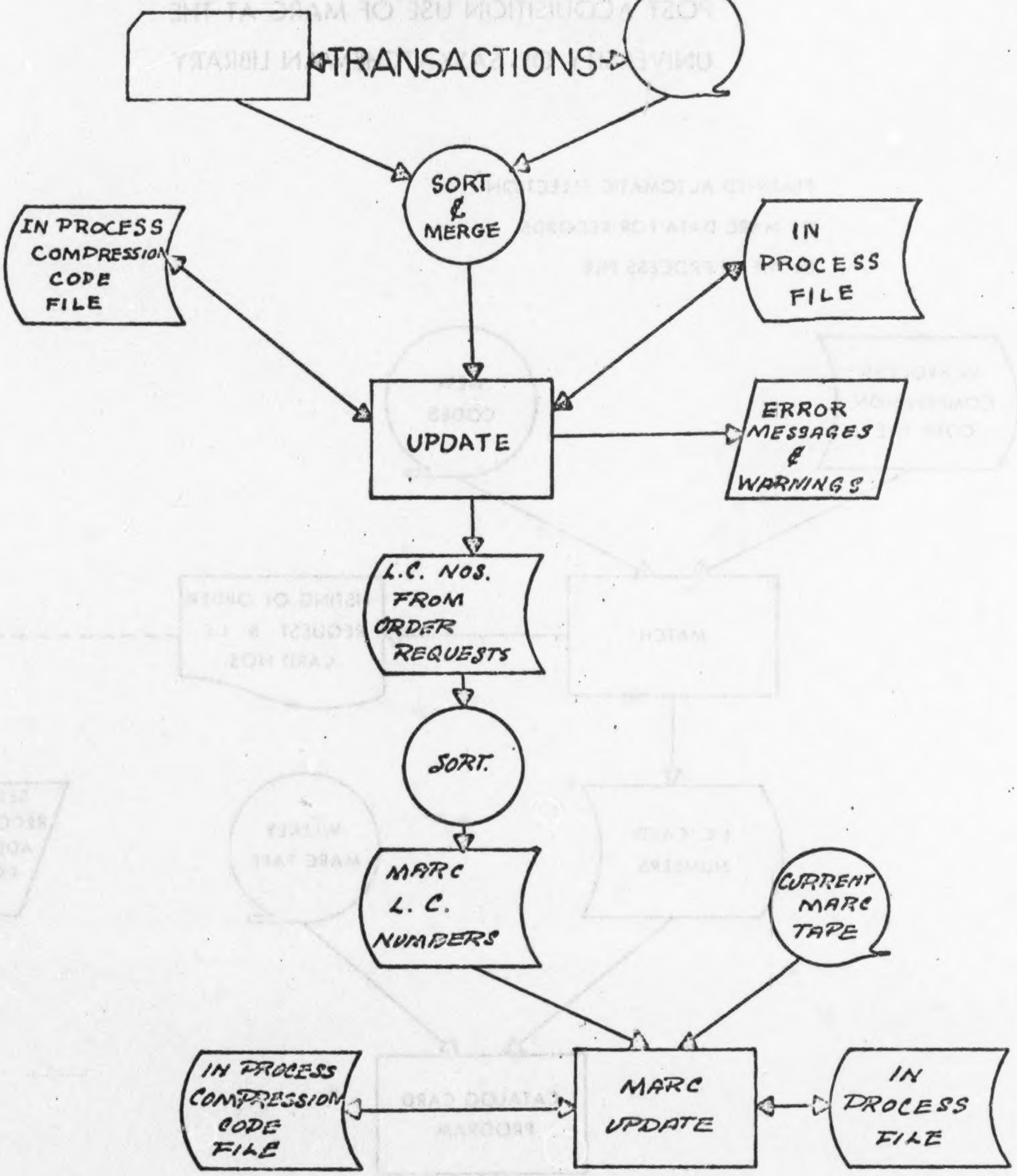
```
                    LIBRARY - ENTRY/TITLE REQUESTS FOR MARC CATALOG CARDS     PAGE 10
           REQUEST                                            SEARCH CODE        LC CARD NUMBER(S)
           -------                                            -----------        -----------------

  HARRISON                                                    XFBKFCPLHARPSN     71458825
  OXFORD BOOK CF FCOC PLANTS

                                                              XFSTFRFF           70466105    75455729    75460116
  OXFORD STUDIES AFRICAN AFFAIRS                                                 76427016

  REICHENBACH                                                 XMTHRLATREICCH     68021540
  AXIOMATIZATION OF THE THEORY OF RELATIVITY

  ORNSTEIN                                                    XPTME   ORNSTN   NO LC CARD NUMBERS FOUND
  ON THE EXPERIENCE CF TIME

  041000196                                                   041000196          71443607

  04120123                                                    04120123           70466C93

  04150027X                                                   04150027X          74460085

  043301455                                                   043301495        NO LC CARD NUMBERS FOUND

  043340032                                                   043340032        NO LC CARD NUMBERS FOUND

  048900028                                                   048900028          74448277

  049310054                                                   049310054          74438205

  049310062                                                   049310062          77440120

  0494208C1                                                   049420801        NO LC CARD NUMBERS FOUND

  09C988019                                                   090988019        NO LC CARD NUMBERS FOUND

  0945164C5                                                   094516405          77462928    78474560

  155359802                                                   155359802        NO LC CARD NUMBERS FOUND

  224616579                                                   224616579          72476309

  224617885                                                   224617885        NO LC CARD NUMBERS FOUND
```

POST ACQUISITION USE OF MARC AT THE UNIVERSITY OF SASKATCHEWAN LIBRARY

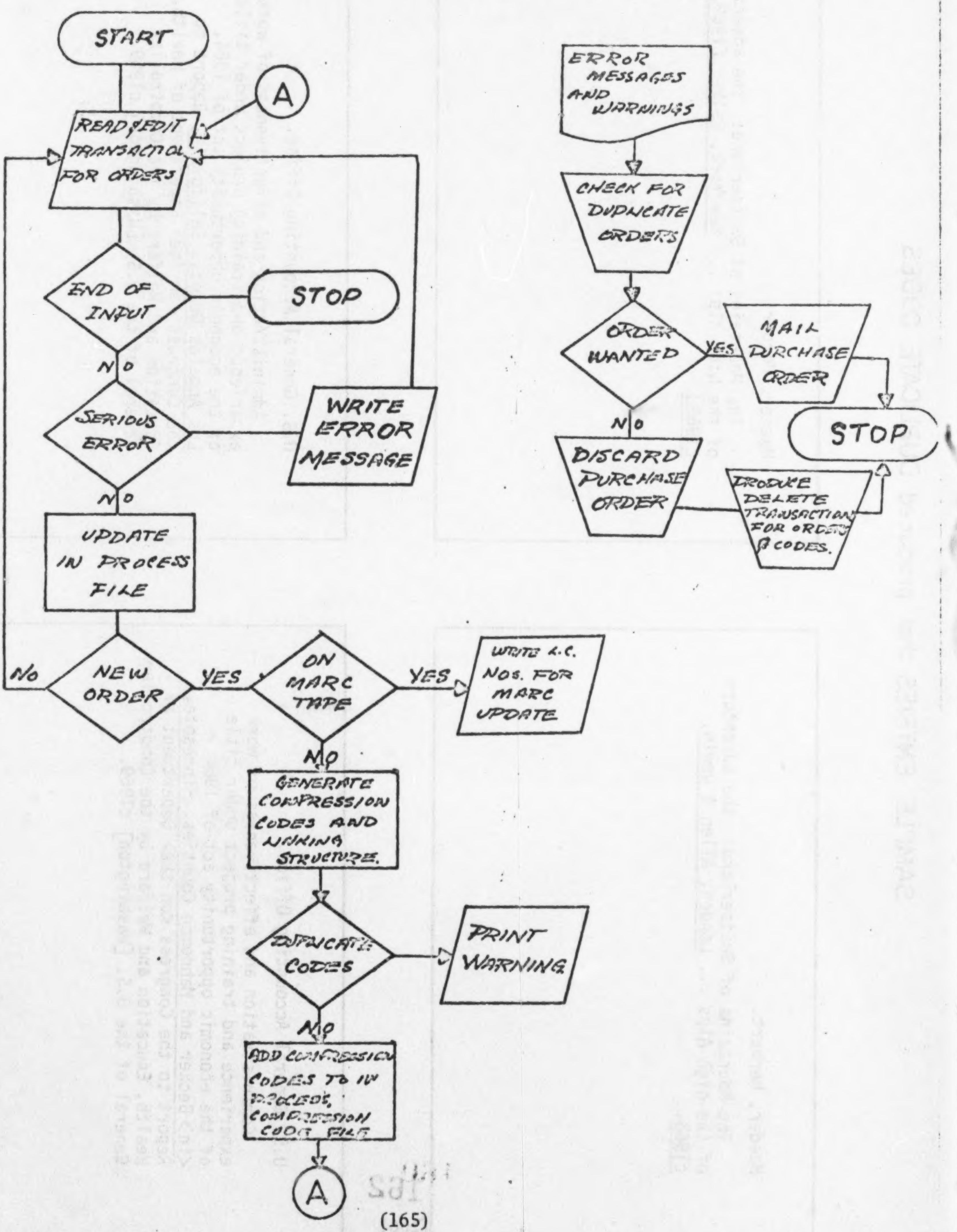
PLANNED AUTOMATIC SELECTION
OF MARC DATA FOR RECORDS
IN THE IN PROCESS FILE


IN PROCESS COMPRESSION CODE FILE
NEW CODES
MATCH
LISTING OF ORDER REQUEST & L.C. CARD NOS.
L.C. CARD NUMBERS
WEEKLY MARC TAPE
SELECT MARC RECORDS TO BE ADDED TO IN PROGRESS FILE
CATALOG CARD PROGRAM
CATALOG CARDS

MARC INPUT IN TO TESA 1
U. of S. LIBRARY


TRANSACTIONS
SORT & MERGE
IN PROCESS COMPRESSION CODE FILE
IN PROCESS FILE
UPDATE
ERROR MESSAGES & WARNINGS
L.C. NOS. FROM ORDER REQUESTS
SORT.
MARC L.C. NUMBERS
CURRENT MARC TAPE
IN PROCESS COMPRESSION CODE FILE
MARC UPDATE
IN PROCESS FILE

TESA 1 UPDATE U. of S. LIBRARY


START
A
READ & EDIT TRANSACTION FOR ORDERS
END OF INPUT
STOP
NO
SERIOUS ERROR
WRITE ERROR MESSAGE
NO
UPDATE IN PROCESS FILE
NO
NEW ORDER
YES
ON MARC TAPE
YES
WRITE L.C. NOS. FOR MARC UPDATE
NO
GENERATE COMPRESSION CODES AND LINKING STRUCTURE.
DUPLICATE CODES
PRINT WARNING
NO
ADD COMPRESSION CODES TO IN PROCESS, COMPRESSION CODE FILE
A
ERROR MESSAGES AND WARNINGS
CHECK FOR DUPLICATE ORDERS
ORDER WANTED
YES
MAIL PURCHASE ORDER
STOP
NO
DISCARD PURCHASE ORDER
PRODUCE DELETE TRANSACTION FOR ORDERS & CODES.

SAMPLE ENTRIES that produced DUPLICATE CODES

Maeder, Herbert.

The Mountains of Switzerland: the adventure of the high Alps ... London, Allen & Unwin, c1968.

Maeder, Herbert.

The Mountains of Switzerland: the adventure of the high Alps ... New York, Walker [1969, c1968]

U.S. General Accounting Office.

Administration and effectiveness of work experience and training project under title of the economic opportunity act of 1964 <in> Becker and Mahnomen Counties, Minnesota, Report to the Congress <on the> Department of Health, Education and Welfare by the Comptroller General of the U.S. [Washington] c1969.

U.S. General Accounting Office.

Administration and effectiveness of work experience and training project under title of the economic opportunity act of 1964, Los Angeles County, California, Report to the Congress <on the> Department of Health, Education and Welfare by the Comptroller General of the U.S. [Washington] c1969.

MARC Costs to date
University of Saskatchewan Library

Month	Development Cost	Update Cost	MARC printout Costs	No. of Printouts
April '69	$ 664.05	$ 4.50		
May	380.49	120.10	$ 27.10	
June	453.01	135.80	79.58	
July	343.03	108.77	44.70	166
August	368.76	150.20	81.70	185
September	439.20	134.30	122.20	111
October	131.63	194.60	153.00	388
November	253.60	197.60	175.70	371
December	105.53	243.90	141.00	491
January '70	409.38	271.20	113.10	489
February	276.00	233.60	200.90	585
March	163.99	189.90	41.93	376
April	72.56	164.81	54.47	747
Totals	$ 4,061.23	$ 2,149.28	$ 1,235.38	3909

TOTAL MARC Costs $ 7,445.89

Title: The War of 1812.

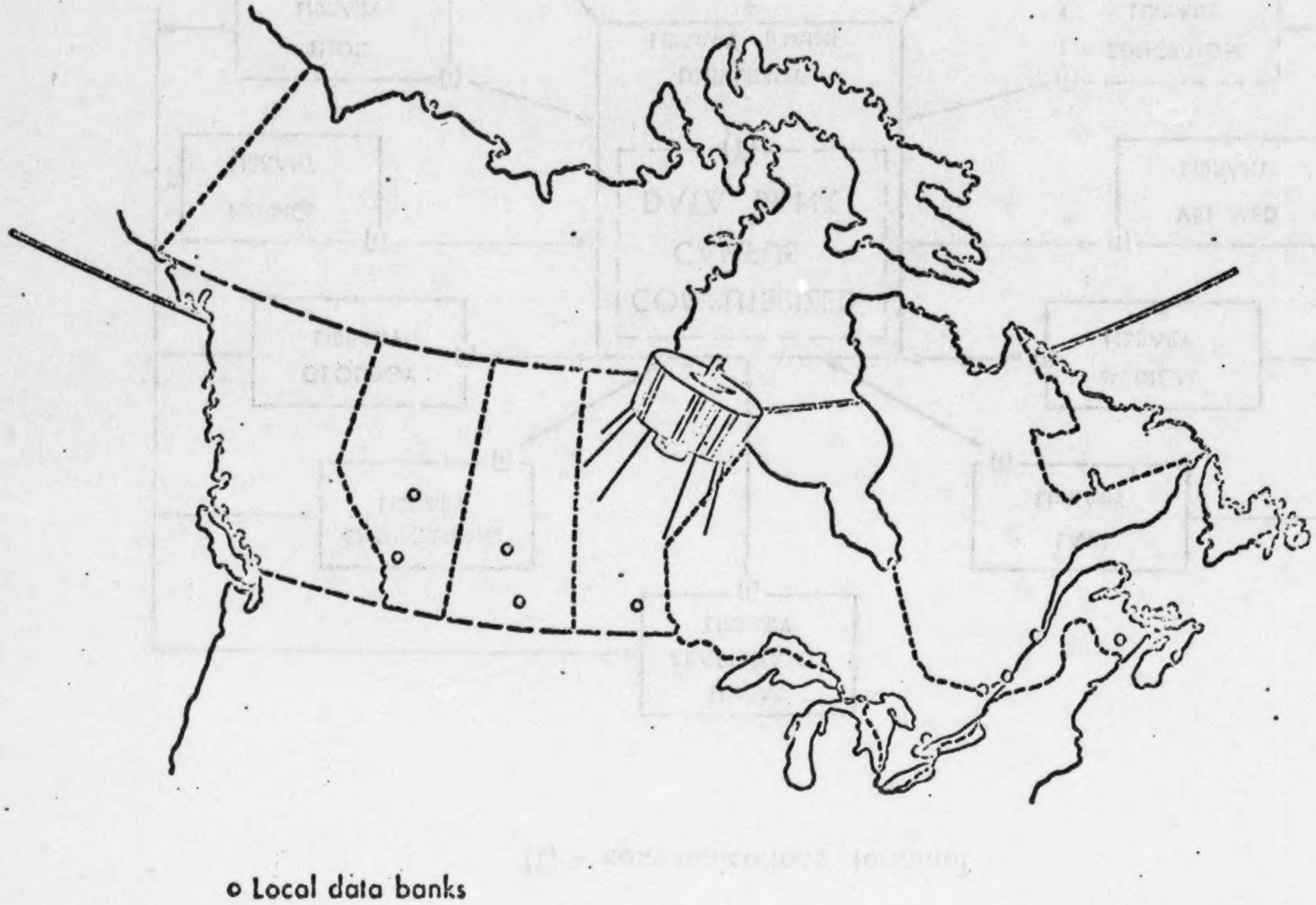
Should print as: The war of 1812.

Should sort as: War of eighteen twelve

This will be accomplished by carrying the field in the record as follows:

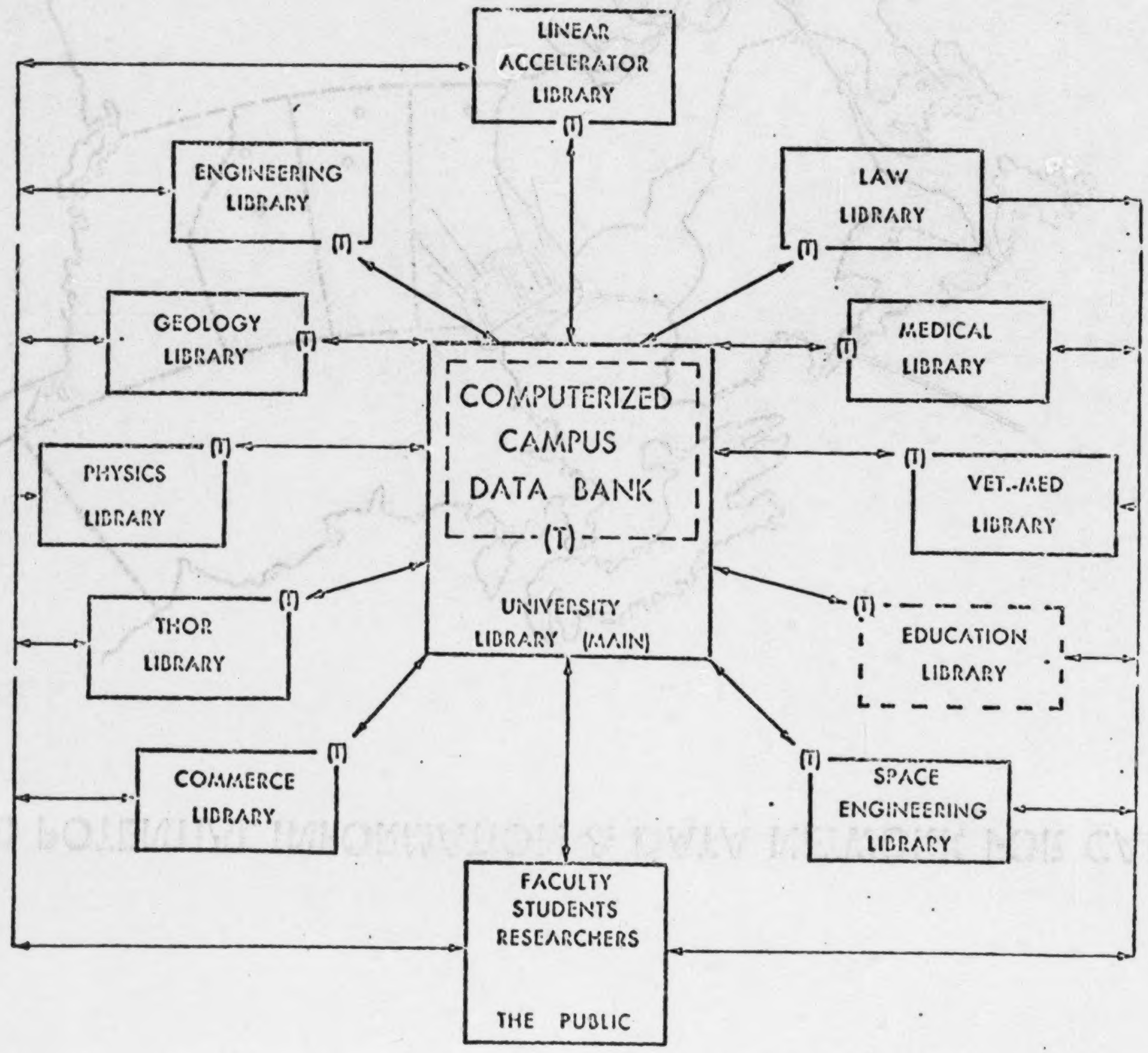
\The \\War of\1812.\eighteen twelve|

A SIMPLIFIED POTENTIAL INFORMATION & DATA NETWORK FOR CANADA


o Local data banks

POTENTIAL
U of S LIBRARY NETWORK

(T) = communications terminal


LINEAR ACCELERATOR LIBRARY
(T)
ENGINEERING LIBRARY
(T)
LAW LIBRARY
(T)
GEOLOGY LIBRARY
(T)
MEDICAL LIBRARY
(T)
COMPUTERIZED CAMPUS DATA BANK
(T)
UNIVERSITY LIBRARY (MAIN)
PHYSICS LIBRARY
(T)
VET.-MED LIBRARY
(T)
THOR LIBRARY
(T)
EDUCATION LIBRARY
(T)
COMMERCE LIBRARY
(T)
SPACE ENGINEERING LIBRARY
(T)
FACULTY STUDENTS RESEARCHERS
THE PUBLIC

Panel Discussion On The Joys and Rewards, or Trials and Tribulations, of Automating a Library

T.C. Dobb,
Chairman

Panel Discussion On The Joys and Rewards, or Trials and Tribulations, of Automating a Library

T.C. Dobb, Chairman

T.C. DOBB: Introductory Remarks

We have taken some liberties with the topic. We are going to talk about systems analysis because it's a very important subject still misunderstood by a great number of people.

Librarians didn't start talking about systems analysis until they started talking about automation. And that's too bad because now when you mention systems analysis to them, they assume you mean to start automating something.

It's true you can't automate without systems analysis coming in somewhere -- usually *before* the programming. Although on a recent project at S.F.U. we tried -- with more audacity than intelligence -- to do the systems work *after* the programming. *Don't* try it. It doesn't work. That's a touch of tribulation for you.

But if you're not automating any part of your Library's procedures, you're probably doing one of two things as far as systems analysis is concerned:

1. You're not doing any analysis at all. In which case you're running a bad operation. Meaning that, if you had to declare a profit, you'd be bankrupt in the first six months.
2. You're doing analysis accidentally because some of your people have tidy minds. For example:

A few years ago in the Acquisitions Section at S.F.U. we had shelves of books waiting for L.C. cards. Each one of these books had a deck of punched cards in it. The decks were always falling out of the books or being bent by the bottom of the shelf above. One day we borrowed a girl from another part of the Library to help us match L.C. cards with books. She wasn't a librarian but she was sexy and she had a tidy mind. Well she worked with us for a couple of hours, then suggested that we keep all those books on their spines. An elegant solution: the decks no longer fell out nor were they bent by the shelf above. In addition, the purchase order numbers could be read without breaking your neck. That's a touch of joy for you.

A problem as small as that doesn't require the formal methodology of systems analysis. But even if you have a library full of sexy women with tidy minds, they won't be able to solve your large systems problems without resorting to that formal methodology.

What Mel and Don will say in the next few minutes will be directed at encouraging you to believe that this is the case.

M. ENDLEMAN:

If you were asked to define the outstanding traits of a System Analyst, you will probably set the response "He understands how a computer works and ..." a long pause.

Too little attention is given to the qualities, other than technical, that make up a System Analyst.

The System Analyst recognizes that automation (be it a computer or a conveyor belt) is only one phase or element in the project. The Analyst, because he is often the only individual who has access to the total picture, must relate and assume responsibility for both the manual and automated phases. He must apportion the time to each phase so that it will interface completely with the whole project.

The effective Analyst will spend his most concentrated effort on the identification phase. The designing of a system and solving the wrong problem is always a possibility.

The synergistic approach to problem solving shows the better Analyst has confidence in ideas tested in a group environment. This team method also helps catch many of the small details that one person may miss.

The Analyst also has the responsibility to present alternative solutions to the managers and ultimate user. He will also dig deeply to use valid suggestions from the people already doing the job.

The ability to observe and use of the ability to observe and react to the observations in the daily ritual of your job, makes everybody an Analyst. The only problem to be aware of, is, when making an adjustment for your own benefit, find out what that change will do to other departments.

FOR EXAMPLE:

At peak times, the borrowers were lined up throughout much of the main floor, waiting to check out their books. Delays up to 10 minutes at these peak periods were common.

The problem was the borrower did not have space available on the approach to the check out machine to open the books for the girl to pull out the book card. Therefore she had to spend time opening the books as well as pulling out the book card.

Solution was to build a ten foot extension for the borrowers to have space to open their books and also to make this extension double width so two rows of borrowers could be accommodated by having two girls man the check out desk.

So the Analyst is not a magician or an ogre who changes things for the sake of changing. But an observationist with the job of acting on those observations.

The Concept of Systems Analysis an Essential Component of Management

by M. Sanderson, Library

Introduction

Why take so long? I want to do it now!

What is there to plan?

Where's the problem in designing a file?

But this isn't a computer system.

What do you mean by a system?

This article has been written with the intention of providing an insight into the nature of systems analysis such that the kinds of question represented above will no longer take the form of exasperated demands.

The use of a systems group in business and industry has become an essential part of the management process. The need for systems analysts in all large, organized enterprises is rapidly becoming acute. What is it that makes this the case?

It is my purpose here to try and present the viewpoint of a systems analyst and his attitudes and techniques in the approach to a problem; also to present the case for a systems approach as a general method of tackling any problem.

This is not intended to be a definitive treatise on systems analysis - there are plenty of how-to-do-it books already available.* What I am concerned with is encouraging an attitude so that these books are read, appreciated and their principles put into action.

---

* If you wish to know how we do it, consult the Library Systems Division's Systems Documentation and Procedures Manual L/SD/6 and "Project Control and the MRDB Project" L/SD/16.

Systems Analysis, Prefatory Remarks

Systems analysis is a systematic way of solving problems and forming a sound framework within which to make decisions. Its method is a detailed examination of a problem area. The problem is viewed in relation to the whole field of which it normally forms a segment. Possible solutions are examined on the basis of their cost in relation to the associated benefits and the ramifications of the changes resulting.

A great deal of personal contact with managerial and administrative personnel is required to provide the information and co-operation necessary to plan and execute projects.

A system is a set of procedures designed to achieve a specific goal. It may be a small part of a total operation; or the sets of procedures which compose it may in themselves be small systems.

A system is thus not just something that concerns computer personnel. Any person working in an organization is in some way attempting to contribute to the achievement of the organization's goals. In this respect he is involved in a system by the operations he performs each day.

A systems group will try to help people clarify their objectives and design operations which are efficient in the sense that work flow is smooth, with no delays, waste of effort or money. General aims are: 1) to provide better planning and decision making information and to furnish more timely and effective reports on operations. 2) To promote the efficient and economical use of man-power, manual and data-processing equipment, communications facilities and money. 3) To simplify and standardize procedures, and promote written policy and procedures manuals to provide continuity of operations in the event of personnel changes. 4) To keep abreast of advances in techniques and equipment.

Systems people should aim at providing a co-ordinated approach to the operations of an outfit as a whole; to look at things from an overall viewpoint rather than deal in isolated fragments. It's all too easy to make improvements conflict with the general aims of the whole operation.

The Systems Concept and Management

The process of management is a complex one which requires the ability to operate in two seemingly contradictory ways: namely to make quick intuitive decisions to solve immediate problems, and logically and systematically to plan and accomplish long-range objectives. Most managers seem to be either one or the other.

There seems to be some general agreement in the literature as to the functions of a manager. These appear to form a cycle composed of the following activities:

1) to determine policy and objectives
2) to plan and to organize effort
3) to control progress
4) to solve problems

The activities of management occur, as mentioned above, at two different levels: firstly, the long-term planning and consideration of overall objectives and secondly the "fire-fighting" approach of making quick decisions in critical situations. However, these need not be viewed as contradictory approaches. The intuitive approach based on experienced judgement, and the systematic approach should be complementary.

One activity which all the above processes have in common to a greater or lesser degree is decision-making.

In the past few years increasing emphasis has been placed on operations research, management science and systems analysis as aids to effective management decision making. The systems approach is basic to decision making processes and emphasizes the consideration of problems as a part of the total complex, rather than as isolated phenomena.

A principal requirement in the solving of problems is the ready availability of appropriate information. This entails the development of good management information systems. Consideration of the information channels which provide the data for decision-making is an integral part of any systems investigation Thus, whereas an essential ingredient in good management is to be able to acquit oneself well in the "quick and dirty", intuitive solutions required to allow the organization to continue to pursue its goals, this must be complemented by a thorough understanding of and application of decision and problem-solving theory and information systems. As Managers become more analytical in their approach to decision making and the nature of problems in large organizations becomes more complex, particularly with the proliferation of automation and communications equipment, they need help in choosing the best strategy from those available.

Many problem situtations overlap divisions and departments and are not clearly attributable to one cause, nor fall to the lot of any particular manager or department head.

The old hierarchical arrangement of an organization may often militate against the effective accomplishment of a project unless a project team is set up to cut across traditional boundaries and reach conclusions not biased towards the needs of one particular department. The manager can no longer survive if he clings to old parochial ideas regarding his own department.

The systems approach - whether a project team formed of analysts and members of a number of involved departments, or a lone team of analysts - is to view the entire system and solve the problem accommodating the objectives of different functional units.

Systems analysis is not required for every decision-making situation. The essential component of some problems is speed of action and the deliberations of systems analysis should not be forced upon these. However, where there is great uncertainty, many alternatives and large resources to be employed, the systems approach is clearly called for.

Systems Analysis: Techniques and Problems

The general approach of systems analysis is as follows:

1) Definition of the problems and objectives
2) Research of fact gathering
3) Search for possible solutions
4) Systematic evaluation of alternatives
5) Implementation of preferred solution
6) System follow-up

Often these processes are arranged in distinct phases for convenience's sake.

1. Feasibility Study

This is an essential preliminary. It involves defining the problem and clarifying objectives. At this point it is decided whether or not a project is a worthwhile undertaking; i.e. would it cost too much, would the disadvantages outweigh the advantages, are the personnel available capable of performing the envisaged operations.

2. System Study

If the project seems reasonable, a plan of attack is drawn up. Procedures are examined in detail, surveys and interviews conducted, and all data carefully documented. Generally a report is produced at this point summarizing the existing situation. The aim at this stage is not so much to propose solutions as to learn what will confront the system.

3. System Design

The next thing is to propose various solutions to the problem and to investigate the consequences. The most suitable solution is chosen. All solutions are drawn up in some detail. The whole is fully documented, including the reasons for the choice of a particular solution. A report on the proposed solution, generally including an implementation plan, is produced.

4. Implementation and Follow-up

The implementation of the new system is carefully supervised with special attention paid to availability of documentation, procedures manuals and staff familiarization and training.

The system in operation is observed closely to ensure that the anticipated results are achieved. The project is not considered complete until these results have been achieved.

Some Problematical Areas

Requirements of an organization change and develop. Systems and procedures must be designed in such a way as to allow change with as little disruption as possible. Often systems are patched up as changes occur and after a time the patches become a major component resulting in unwieldly and inefficient juggernauts.

PROBLEM DEFINITION. Frequently the aspect of problem definition is the most difficult part of a system study. Care must be taken to separate the symptoms from the problem.

Incomplete problem definition may be minimized by having the user define his requirements. The analyst is trying to help the user department achieve its goals, so the user should as far as possible define the problem and what output is expected. For the analyst to attempt to define these often leads to peculiar results. It is easy for a person from outside to get a false picture of what a department is aiming at.

Until the dimensions of the problem have been adequately determined it is practically impossible to design a system with any confidence in a satisfactory outcome.

Since the approach to the study conditions the anticipated outcome, the objectives must be determined prior to the start of any system design. Thus it is important to sort out exactly what is wanted, and how it is expected that this will be achieved.

Communication and Cooperation

Cooperation of user departments can sometimes be a problem. This is sometimes aggravated by the attitude of the analyst. It is difficult enough for a person who is trying to do his job while being asked all kinds of questions as to how, why, where, when, who. It is even more difficult if the analyst does not explain why he's asking questions, or if he talks down to those involved, or hides behind a screen of jargon.

Also people are not necessarily adjusted to the concepts of systems or automation and don't think in terms of disc space and time-slicing or even cost-benefit analysis or trend.

Nor do analysts always consider the needs of the user carefully. They may design forms, for example, that are either awkward to fill out, or use codes that have to be figured out, or have other defects from the human point of view.

The systems analyst often has difficulty determining what information is available. Frequently the person most familiar with the current system cannot specify exceptions which are now resolved by human judgement. The more exacting input requirements of a computer application are not always easily understood by users familiar with a more flexible, less precise manual system.

Management Involvement

Managerial levels often don't involve themselves to the extent necessary to ensure that projects fit into overall plans and that the appropriate policy decisions are made - or that authority is given to those doing the job. Sometimes management personnel are working full time on other assignments and do not have time to make thorough studies, perhaps leaving development of procedures to insufficiently qualified or trained people.

Scheduling

The things that are detrimental to otherwise good systems are i) no design freeze - the reluctance of the user's department to stop changing their minds and requesting alterations while systems design is in process. ii) failure to arrange proper scheduling of input to a system, whether manual or automated, resulting in overload or lost information and creeping chaos.

Documentation

Documented procedures are in themselves an aid to planning, apart from being vital to efficient operation, since they give a quick picture of the current system. But too often the only information on procedures is stored in certain people's heads. If documentation exists at all it often produces some subtle torments by being out of data.

If documentation is not available, good systems can decay because the people who knew the ins and outs have left. New people take over, don't understand what is going on, and resort to more primitive methods based on expediency.

Systems Analysis and S.F.U. Library

The principal objective of the Library Systems Division is to promote efficient systems in the Library. It is intended that this function will be enterprising and continuous, not static or intermittent, with emphasis on development, and continually seeking new and better approaches to performing activities.

This will be a cooperative effort between the Systems Division and the professional librarians. The Systems Division will provide the systems and analytical tools and automation methods. The Librarians will provide direction in terms of library aims, and advice in terms of specialized knowledge.

The Systems Division will also be responsible for liaison between the Library and the Computing Centre.

Some people in the computer world feel that systems designed for computer applications will only do what are often very trivial things in highly effective ways until management and people involved with them can be educated to the possibilities of computer systems. Part of the task of the Systems group is to make people aware that computers can do things other than provide long lists of unnecessary information which no one reads; also to explain to people that it is not justifiable to put on-line information that is perfectly adequate on a 24-hour turnaround basis.

Conclusion

The systems approach is not restricted to automated systems. Because a system will not be subject to the rigorous demands of a computer application does not mean that we may lapse into woolly thinking and badly thought our procedures. When given thought it can be seen that the demands of sophisticated management are equally rigorous. The old computer dictum "Garbage in, garbage out" is perfectly applicable to the management planning function.

Given the acceptance in an organization of, and familiarity with, the systems approach, the systems group can spend less of its time on aid in revision of existing minor procedures, and resisting wild innovations, and turn its effort to systems development and the study of the total system.